STIC-Biotech/ChemLib

76379

From: Fredman, Jeffrey
Sent: Tuesday, September 24, 2002 10:20 AM
To: STIC-Biotech/ChemLib
Cc: Schmidt, Mary
Subject: FW: sequence search request 09/923,831

I Approve.

Please Rush.

Jeff Fredman

-----Original Message-----
From: Schmidt, Mary
Sent: Tuesday, September 24, 2002 9:53 AM
T : Fredman, Jeffrey
Subject: RE: sequence search request 09/923,831

2365 bases

-----Original Message-----
Fr m: Fredman, Jeffrey
Sent: Tuesday, September 24, 2002 7:06 AM
To: Schmidt, Mary
Subject: RE: sequence search request 09/923,831

How big is the sequence?

-----Original Message-----
From: Schmidt, Mary
Sent: Monday, September 23, 2002 3:33 PM
To: Fredman, Jeffrey
Subject: FW: sequence search request 09/923,831

Hi, I ordered a full length search of SEQ ID NO:42 last month and got the results, however, I didn't notice then that they also have fragment claims-- could you please rush the following request?

Please search SEQ ID NO:42: (1) results size limited to less than 100 bases, and (2) results size limited from 100-300 bases. (NO interference search and NO full-length search).
Thanks!!

Melissa
11D05
mailbox 11E12
au. 1635

We can only limit and not range. So, there is 100 & 300 only

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/24/02
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 0.25
WWW/Internet: _____
Other (specify): _____

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into to two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:15:48 ; Search time 4266.96 Seconds
(without alignments)
11598.716 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 1365494

Minimum DB seq length: 0
Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	62	2.6	234	6	AX314960	AX314960 Sequence
	2	59.8	2.5	229	11	G66079	G66079 sY2139 YAC
	3	50.8	2.1	235	1	STREXP9B	L20564 Streptococc
	4	50.8	2.1	235	6	AR084893	AR084893 Sequence
c	5	47	2.0	224	3	DDIACTD	M29112 D.discoideu
	6	47	2.0	276	6	AX308194	AX308194 Sequence
c	7	46.8	2.0	247	6	AX186933	AX186933 Sequence
	8	46.8	2.0	280	8	AB027882	AB027882 Schizosac
c	9	46.4	2.0	268	6	AX185659	AX185659 Sequence
c	10	44.4	1.9	220	6	AX185507	AX185507 Sequence
c	11	44.4	1.9	300	6	AX185628	AX185628 Sequence
c	12	44	1.9	160	3	DDIACTE	M29113 D.discoideu
c	13	43.6	1.8	294	6	AX185661	AX185661 Sequence
	14	43.4	1.8	140	3	DMRRNL	X08015 Drosophila
c	15	43	1.8	288	6	AX185400	AX185400 Sequence
c	16	42.4	1.8	220	11	G58490	G58490 SHGC-104454
c	17	42.2	1.8	174	6	AX340337	AX340337 Sequence
c	18	41.8	1.8	289	6	AX198818	AX198818 Sequence
	19	41.6	1.8	229	6	AX062568	AX062568 Sequence
c	20	41.4	1.8	255	6	AX210942	AX210942 Sequence
	21	41.4	1.8	272	3	ECA270206	AJ270206 Entodiniu
c	22	41.2	1.7	255	6	AX211073	AX211073 Sequence
	23	41	1.7	116	6	A08900	A08900 H.sapiens (
	24	41	1.7	119	6	A08901	A08901 H.sapiens (
	25	41	1.7	119	6	A08904	A08904 H.sapiens (
c	26	41	1.7	198	6	AX182091	AX182091 Sequence
	27	41	1.7	202	6	AX099441	AX099441 Sequence
	28	41	1.7	206	10	AF168586	AF168586 Rattus no
c	29	40.8	1.7	250	6	AX210706	AX210706 Sequence
	30	40.6	1.7	192	11	G38017	G38017 C1M12 Plasm
	31	40.6	1.7	263	11	CNS06EHA	AL395108 T3 end of
c	32	40.4	1.7	226	6	AX184991	AX184991 Sequence
	33	40.2	1.7	190	11	HUMUT5193A	L31113 Human STS U
	34	40.2	1.7	252	13	AF303464	AF303464 Unidentif
	35	40	1.7	143	6	AX255922	AX255922 Sequence
c	36	39.8	1.7	233	6	AX182078	AX182078 Sequence
	37	39.8	1.7	240	8	MITGIR1	X02170 Torulopsis
c	38	39.6	1.7	112	6	AX261671	AX261671 Sequence
	39	39.6	1.7	202	4	DOGSNVD17B	M73046 Dog inserte
	40	39.4	1.7	99	6	A08898	A08898 H.sapiens (
	41	39.4	1.7	104	6	A08899	A08899 H.sapiens (
	42	39.4	1.7	222	6	AR122013	AR122013 Sequence
c	43	39.4	1.7	252	6	AX197621	AX197621 Sequence
	44	39.4	1.7	258	10	BC014856	BC014856 Mus muscu
	45	39.4	1.7	295	6	AX184943	AX184943 Sequence

ALIGNMENTS

RESULT 1
AX314960
LOCUS AX314960 234 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 7945 from Patent WO0190366.
ACCESSION AX314960
VERSION AX314960.1 GI:17898715
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Leach,M.D. and Shimkets,R.A.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0190366-A 7945 29-NOV-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1. .234
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 59 a 60 c 54 g 61 t
ORIGIN

```
  Query Match             2.6%;  Score 62;  DB 6;  Length 234;
  Best Local Similarity   56.3%;  Pred. No. 0.0051;
  Matches  116;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy   1757 tgcatggagatagagaacagagagatcgggagaaagcattagagaactttaaaacaggca 1816
          | ||||| || |    ||    ||  | ||    || || |    ||| | | | ||
Db     29 TTCATGGTGACCGTACCCAATCTGAGAGAGAACGTGCCTTGGCCGCCTTCAGATCTGGTG 88

Qy   1817 aagtgagaatactaattgcaactgatctagcctctagaggacttgatgtccatgacgtta 1876
          |  |   || |  |||| || | | | ||  ||||||| || ||| | |   |||| |
Db     89 CCGCTACTTTATTGGTTGCCACAGCTGTCGCAGCTAGAGGTCTAGATATTCCAAACGTCA 148

Qy   1877 cacatgtctataattttgactttccacggaatattgaagaatacgtacaccgaatagggc 1936
          | || ||    || |  || || ||| |  || | || || ||||| ||  |||| ||
Db    149 CCCACGTTATCAACTACGATTTACCAAGTGATGTCGATGATTACGTCCATAGAATTGGTA 208

Qy   1937 gcacgggaagagcagggaggactggt 1962
          | || ||  | || || |  || |||
Db    209 GAACTGGTCGTGCCGGTAACACCGGT 234


RESULT  2
G66079
LOCUS       G66079                   229 bp    DNA     linear   STS 02-NOV-2001
DEFINITION  sY2139 YAC Subtraction Homo sapiens STS genomic, sequence tagged
            site.
ACCESSION   G66079
VERSION     G66079.1  GI:15078280
KEYWORDS    STS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 229)
  AUTHORS   Tilford,C.A., Kuroda-Kawaguchi,T., Skaletsky,H., Rozen,S.,
            Brown,L.G., Rosenberg,M., McPherson,J.D., Wylie,K., Sekhon,M.,
            Kucaba,T.A., Waterston,R.H. and Page,D.C.
  TITLE     A physical map of the human Y chromosome
  JOURNAL   Nature 409 (6822), 943-945 (2001)
  MEDLINE   21131744
COMMENT     Synonyms: SP10-22
            Contact: Charles A. Tilford
            Page Lab
            Whitehead Institute for Biomedical Research
            Room 425, 9 Cambridge Center, Cambridge, MA 02142, USA
            Email: tilford@wi.mit.edu
            Primer A: TGAGGGAGTTTCATTCTGGC
            Primer B: AACCTGCTGTACATCAATGCC
            STS size: 80
            PCR Profile:
                 94C        4:00 min
                   / 94C     :30 sec
                 35x | 56C   2:00 min
                   \ 72C     2:00 min
                 72C        7:00 min
            Protocol:
                 Template:       25-100 ng
                 Primer:         each 1 uM
                 dNTPs:          each 100 uM
                 Taq Polymerase: 0.05 units/ul
                 Total Vol:      15 ul
                 Mineral Oil:    15 ul

            Buffer:
                For 20ml of 10x Stock solution:
                Stock Reagent          Volume Mixed    Final Conc. of 10x Stock
                1M KCl                   10 ml          500 mM
                1M Tris-HCl, pH 9.0       2 ml          100 mM
                Triton X-100            200 ul          1.0 %
                2M MgCl2                150 ul           15 mM
                H2O                    7.65 ml
```

```
            The base sequence listed above was later found to contain errors.
            Although these primers still constitute a functional STS, we
            believe at this time that the following sequence more accurately
            reflects the base sequnce:
            GATCATGAGGGAGTTTCATTCTGGCTCTAGCAGAGTTTTGATTACAACTGACCTGCTGGC
            CAGAGGCATTGATGTGCAGCAGGTTTCTTTAGTCATCAACTATGACCTTCCCACCAACAG
            GGAAAACTATATCCACAGAATCAGTCGAGGTGGACGGTTTGGCCGTAAAGGTGTGGCTAT
            TAACATGGTGACAGAAGAAGACAAGAGGACTCTTCAAGACATTGAGATC.
FEATURES             Location/Qualifiers
     source          1. .229
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /sex="male"
                     /cell_line="OXEN XYYYY Male"
                     /clone_lib="YAC Subtraction"
                     /note="Vector: pYAC4; V-type: YAC; DNA subtraction between
                     Y chromosome YACs used to generate a pool of  fragments
                     enriched for low-copy number, Y-specific sequences."
     STS             6. .85
     primer_bind     6. .25
     primer_bind     complement(65. .85)
BASE COUNT       67 a     47 c     58 g     57 t
ORIGIN


  Query Match             2.5%;  Score 59.8;  DB 11;  Length 229;
  Best Local Similarity   56.3%;  Pred. No. 0.015;
  Matches  112;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy   1800 gaactttaaaacaggcaaagtgagaatactaattgcaactgatctagcctctagaggact 1859
          | | ||| |  | |||      ||| |  | ||| |||||||  ||     | |||||  |
Db     10 GGAGTTTCATTCTGGCTCTAGCAGAGTTTTGATTACAACTGACCTGCTGGCCAGAGGCAT 69

Qy   1860 tgatgtccatgacgttacacatgtctataattttgactttccacggaatattgaagaata 1919
          ||||||  ||  | ||| |    |||   || | |||| ||||    || |  ||| | ||
Db     70 TGATGTACAGCAGGTTTCTTTAGTCATCAACTATGACCTTCCCACCAACAGGGAAAACTA 129

Qy   1920 cgtacaccgaatagggcgcacgggaagagcagggaggactggtgtttccattacaacttt 1979
            | ||| ||||  | ||    ||| |    ||  | |  |||||  | ||||  |   |
Db    130 TATCCACAGAATCAGTCGAGGTGGACGGTTTGGCCGCAAAGGTGTGGCTATTAACATGGT 189

Qy   1980 gactagaaatgattggagg 1998
          |||   | | ||   ||||
Db    190 GACAGAAGAAGACAAGAGG 208


RESULT  3
STREXP9B
LOCUS       STREXP9B                 235 bp    DNA     linear   BCT 14-OCT-1993
DEFINITION  Streptococcus pneumoniae Exp9 gene, partial cds.
ACCESSION   L20564
VERSION     L20564.1  GI:398102
KEYWORDS    Exp9 protein.
SOURCE      Streptococcus pneumoniae (strain R6x) DNA.
  ORGANISM  Streptococcus pneumoniae
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 235)
  AUTHORS   Pearce,B.J., Yin,Y.B. and Masure,H.R.
  TITLE     Genetic identification of exported proteins in Streptococcus
            pneumoniae
  JOURNAL   Mol. Microbiol. 9, 1037-1050 (1993)
  MEDLINE   95020625
FEATURES             Location/Qualifiers
     source          1. .235
                     /organism="Streptococcus pneumoniae"
                     /strain="R6x"
                     /db_xref="taxon:1313"
     CDS             <1. .>234
                     /codon_start=1
                     /transl_table=11
                     /product="Exp9"
                     /protein_id="AAA26886.1"
```

```
                     /db_xref="GI:398103"
                     /translation="AFVFGRTKRRVDELTRGLKIRGFRAEGIHGDLDQNKRLRVLRDF
                     KNGNLDVLVATDVAARGLDISGVTHVYNYDIPQD"
BASE COUNT       58 a     48 c     58 g     71 t
ORIGIN


  Query Match             2.1%;  Score 50.8;  DB 1;  Length 235;
  Best Local Similarity   51.3%;  Pred. No. 1.3;
  Matches  118;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy   1679 ttgtcttcgtttctcgaaaagctgttgcggatcacttatcaagtgacctaatacttggaa 1738
          |||| || | |  |   |||     || |||| | ||  |  |||   | | | || |
Db      5 TTGTATTTGGTCGTACCAAACGCCGTGTGGATGAATTGACTCGTGGTTTGAAAATTCGTG 64

Qy   1739 atatatcagtagagtctctgcatggagatagagaacagagagatcgggagaaagcattag 1798
             |    | |||     | ||||| ||   ||| || |   | ||       |   |
Db     65 GCTTCCGTGCAGAAGGAATTCATGGCGACCTAGACCAAAACAAACGTCTTCGTGTCCTTC 124

Qy   1799 agaactttaaaacaggcaaagtgagaatactaattgcaactgatctagcctctagaggac 1858
             |||||||||  |||||  |     |  |  |||| || ||  | ||  |  | ||
Db    125 GTGACTTTAAAAATGGCAATCTTGATGTTTTGGTTGCGACAGACGTTGCAGCGCGTGGTT 184

Qy   1859 ttgatgtccatgacgttacacatgtctataattttgactttccacggaat 1908
          | ||| |    |  || || ||||||||| || |  ||  ||||||   ||
Db    185 TGGATATTTCAGGTGTGACCCATGTCTACAACTACGATATTCCACAAGAT 234


RESULT   4
AR084893
LOCUS       AR084893                 235 bp    DNA     linear   PAT 01-SEP-2000
DEFINITION  Sequence 19 from patent US 5981229.
ACCESSION   AR084893
VERSION     AR084893.1  GI:10011664
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 235)
  AUTHORS   Masure,H.Robert, Pearce,B.J. and Tuomanen,E.
  TITLE     Bacterial exported proteins and acellular vaccines based thereon
  JOURNAL   Patent: US 5981229-A 19 09-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .235
                     /organism="unknown"
BASE COUNT       58 a     48 c     58 g     71 t
ORIGIN


  Query Match             2.1%;  Score 50.8;  DB 6;  Length 235;
  Best Local Similarity   51.3%;  Pred. No. 1.3;
  Matches  118;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy   1679 ttgtcttcgtttctcgaaaagctgttgcggatcacttatcaagtgacctaatacttggaa 1738
          |||| || | |  |   |||     || |||| | ||  |  |||   | | | || |
Db      5 TTGTATTTGGTCGTACCAAACGCCGTGTGGATGAATTGACTCGTGGTTTGAAAATTCGTG 64

Qy   1739 atatatcagtagagtctctgcatggagatagagaacagagagatcgggagaaagcattag 1798
             |    | |||     | ||||| ||   ||| || |   | ||       |   |
Db     65 GCTTCCGTGCAGAAGGAATTCATGGCGACCTAGACCAAAACAAACGTCTTCGTGTCCTTC 124

Qy   1799 agaactttaaaacaggcaaagtgagaatactaattgcaactgatctagcctctagaggac 1858
             |||||||||  |||||  |     |  |  |||| || ||  | ||  |  | ||
Db    125 GTGACTTTAAAAATGGCAATCTTGATGTTTTGGTTGCGACAGACGTTGCAGCGCGTGGTT 184

Qy   1859 ttgatgtccatgacgttacacatgtctataattttgactttccacggaat 1908
          | ||| |    |  || || ||||||||| || |  ||  ||||||   ||
Db    185 TGGATATTTCAGGTGTGACCCATGTCTACAACTACGATATTCCACAAGAT 234


RESULT   5
DDIACTD/c
LOCUS       DDIACTD                  224 bp    DNA     linear   INV 27-APR-1993
DEFINITION  D.discoideum actin A-3-sub-1 gene, 5' flank.
ACCESSION   M29112
VERSION     M29112.1  GI:167619
KEYWORDS    actin.
SOURCE      D.discoideum DNA.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 224)
  AUTHORS   Romans,P. and Firtel,R.A.
  TITLE     Organization of the Dictyostelium discoideum actin multigene
            family: Flanking sequences show subfamily homologies and unusual
            dyad symmetries
  JOURNAL   J. Mol. Biol. 183, 311-326 (1985)
  MEDLINE   85264789
FEATURES             Location/Qualifiers
     source          1. .224
                     /organism="Dictyostelium discoideum"
                     /db_xref="taxon:44689"
BASE COUNT      101 a      4 c      8 g    111 t
ORIGIN


  Query Match             2.0%;  Score 47;  DB 3;  Length 224;
  Best Local Similarity   66.0%;  Pred. No. 8.2;
  Matches   68;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy   2263 tcttaaaataatagtgtttgaaaatatagaatccagtgttttatacttttctttaataaaa 2322
          | | || |||  | |  |||   || |   |     | || | |  ||| || | | |||
Db    191 TTTGAATATATAAATTATTGTTTATTTTTTAAAGTTTTTTATTTTTTTTTTTCATTTAAA 132

Qy   2323 atagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
          | | || ||||||  ||  |||||||||||||||||||||||
Db    131 AAATAATTATTTATACAAAAAAAAAAAAAAAAAAAAAAAAA 89


RESULT   6
AX308194
LOCUS       AX308194                 276 bp    DNA     linear   PAT 14-DEC-2001
DEFINITION  Sequence 1179 from Patent WO0190366.
ACCESSION   AX308194
VERSION     AX308194.1  GI:17895327
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Leach,M.D. and Shimkets,R.A.
  TITLE     Human polynucleotides and polypeptides encoded thereby
  JOURNAL   Patent: WO 0190366-A 1179 29-NOV-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .276
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       52 a     78 c     72 g     74 t
ORIGIN


  Query Match             2.0%;  Score 47;  DB 6;  Length 276;
  Best Local Similarity   54.3%;  Pred. No. 8.3;
  Matches   95;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy   1826 tactaattgcaactgatctagcctctagaggacttgatgtccatgacgttacacatgtct 1885
          | || ||  | || || ||     || | ||  |||||||  ||  | ||  | |  |||
Db     26 TGCTCATCACCACCGACCTGCTTGCTCGTGGTATTGATGTTCAGCAGGTGTCCCTGGTCA 85

Qy   1886 ataattttgactttccacggaatattgaagaatacgtacaccgaatagggcgcacgggaa 1945
             || | |||  | ||     |    ||  | ||| | || || || || ||  |  ||
Db     86 TCAACTATGATCTCCCCACCCAGCCAGAGAATTACCTGCATCGCATTGGTCGTAGTGGTC 145

Qy   1946 gagcagggaggactggtgtttccattacaactttgactagaaatgattggagggt 2000
          |    || ||||  |||||| |||| |    | ||||  |  | ||   | || |
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:56:58 ; Search time 371.69 Seconds
(without alignments)
10924.433 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 2377130

Minimum DB seq length: 0
Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	139	5.9	139	22	AAK24181	Human brain expres
	2	130	5.5	130	22	ABA74723	Human foetal liver
	3	130	5.5	130	22	ABA39434	Probe #17900 for g
	4	130	5.5	130	22	AAK23210	Human brain expres
	5	130	5.5	130	22	AAK49382	Human bone marrow
	6	130	5.5	130	22	AAI26493	Probe #16426 for g
	7	130	5.5	130	22	AAI55243	Probe #23929 used
	8	53.4	2.3	259	21	AAA42361	Human secreted exp
	9	53.4	2.3	299	20	AAX84188	DNA encoding human
	10	53.4	2.3	299	21	AAC79417	5' cDNA sequence o
	11	50.8	2.1	235	16	AAQ83249	Streptococcus pneu
	12	47.2	2.0	204	15	AAQ57421	DEAD ATP helicase
	13	47	2.0	102	24	ABK09532	Human ovarian tumo
c	14	46.8	2.0	197	23	AAS49363	Staphylococcus aur
c	15	46.8	2.0	247	22	AAH71353	Human cervical can
c	16	46.4	2.0	268	22	AAH70080	Human cervical can
c	17	45.2	1.9	282	21	AAF08972	Fusarium venenatum
c	18	44.4	1.9	220	22	AAH69928	Human cervical can
c	19	44.4	1.9	300	22	AAH70049	Human cervical can
c	20	44.2	1.9	165	22	AAI84674	Human polynucleoti
	21	44	1.9	232	22	AAH33958	Human colon cancer
c	22	43.8	1.9	293	22	AAL25205	Human breast cance
	23	43.6	1.8	213	15	AAQ57415	Start factor 4A-II
c	24	43.6	1.8	289	22	AAL16073	Human breast cance
c	25	43.6	1.8	294	22	AAH70082	Human cervical can
	26	43.4	1.8	234	22	AAS29038	cDNA encoding for
	27	43.2	1.8	259	18	AAT91303	Human H1075-1 secr
	28	43	1.8	225	19	AAV26699	Human novel secret
	29	43	1.8	225	21	AAA40538	Human adult ovary
c	30	43	1.8	288	22	AAH69821	Human cervical can
	31	42.6	1.8	191	17	AAT13475	Capture probe for
	32	42.6	1.8	259	18	AAV00423	3' fragment of clo
c	33	42.4	1.8	296	22	AAL16259	Human breast cance
c	34	42.2	1.8	271	22	AAL16362	Human breast cance
	35	42.2	1.8	296	22	AAL09995	Human breast cance
	36	42.2	1.8	296	22	AAL19575	Human breast cance
	37	42	1.8	285	22	AAL34854	Human musculoskele
	38	41.8	1.8	178	22	AAF24361	Retroviral recombi
	39	41.8	1.8	208	22	AAF24363	Retroviral recombi
c	40	41.8	1.8	289	22	AAH83649	Human ovarian tumo
	41	41.8	1.8	295	22	AAS29052	cDNA encoding for
	42	41.6	1.8	229	22	AAF68277	Human lung tumour
	43	41.4	1.8	230	22	AAL00088	Human reproductive
c	44	41.4	1.8	255	22	AAH82075	Rat differential t
	45	41.4	1.8	299	21	AAA00517	Human colon cancer

ALIGNMENTS

RESULT 1
AAK24181/c
ID AAK24181 standard; DNA; 139 BP.
XX
AC AAK24181;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 24172.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 24172; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 139 BP; 34 A; 33 C; 27 G; 45 T; 0 other;

Query Match 5.9%; Score 139; DB 22; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.7e-21;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1814 gcaaagtgagaatactaattgcaactgatctagcctctagaggacttgatgtccatgacg 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCAAAGTGAGAATACTAATTGCAACTGATCTAGCCTCTAGAGGACTTGATGTCCATGACG 80

Qy   1874 ttacacatgtctataattttgactttccacggaatattgaagaatacgtacaccgaatag 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 TTACACATGTCTATAATTTTGACTTTCCACGGAATATTGAAGAATACGTACACCGAATAG 20

Qy   1934 ggcgcacgggaagagcagg 1952
          ||||||||||||||||||||
Db     19 GGCGCACGGGAAGAGCAGG 1
```

RESULT 2
ABA74723
ID ABA74723 standard; DNA; 130 BP.
XX
AC ABA74723;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #23028.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 23028; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;

Query Match 5.5%; Score 130; DB 22; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.4e-19;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy    574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy    634 tgcggaattg 643
          ||||||||||
Db    121 tgcggaattg 130
```

RESULT 3
ABA39434
ID ABA39434 standard; DNA; 130 BP.
XX
AC ABA39434;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #17900 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID No 17900; 530pp; English.

```
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;


  Query Match             5.5%;  Score 130;  DB 22;  Length 130;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-19;
  Matches  130;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy    574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy    634 tgcggaattg 643
          ||||||||||
Db    121 tgcggaattg 130


RESULT   4
AAK23210
ID   AAK23210 standard; DNA; 130 BP.
XX
AC   AAK23210;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 23201.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 23201; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;


  Query Match             5.5%;  Score 130;  DB 22;  Length 130;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-19;
  Matches  130;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy    574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy    634 tgcggaattg 643
          ||||||||||
Db    121 tgcggaattg 130


RESULT   5
AAK49382
ID   AAK49382 standard; DNA; 130 BP.
XX
AC   AAK49382;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 23939.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 23939; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
```

```
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;


  Query Match             5.5%;  Score 130;  DB 22;  Length 130;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-19;
  Matches  130;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy     574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy     634 tgcggaattg 643
           ||||||||||
Db     121 tgcggaattg 130


RESULT   6
AAI26493
ID   AAI26493 standard; DNA; 130 BP.
XX
AC   AAI26493;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #16426 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 16426; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;


  Query Match             5.5%;  Score 130;  DB 22;  Length 130;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-19;
  Matches  130;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy     574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy     634 tgcggaattg 643
           ||||||||||
Db     121 tgcggaattg 130


RESULT   7
AAI55243
ID   AAI55243 standard; DNA; 130 BP.
XX
AC   AAI55243;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #23929 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 23929; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;


  Query Match             5.5%;  Score 130;  DB 22;  Length 130;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-19;
```

```
  Matches  130;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     514 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ataatacaagaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacg 60

Qy     574 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aaagcaaaagcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaa 120

Qy     634 tgcggaattg 643
           ||||||||||
Db     121 tgcggaattg 130


RESULT   8
AAA42361
ID   AAA42361 standard; cDNA; 259 BP.
XX
AC   AAA42361;
XX
DT   21-AUG-2000  (first entry)
XX
DE   Human secreted expressed sequence tag SEQ ID NO:1101.
XX
KW   Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;
KW   expressed sequence tag; EST; probe; chemotactic; proliferative;
KW   immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW   thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW   antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW   antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW   cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW   vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW   insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW   lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW   central nervous system disorder; Alzheimer's disease; stroke;
KW   Parkinson's disease; Huntington's disease; coagulation disorder;
KW   haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW   tumour; infection; depression; psoriasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200021990-A1.
XX
PD   20-APR-2000.
XX
PF   15-OCT-1999;   99WO-US24205.
XX
PR   15-OCT-1998;   98US-0104435.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  McCoy JM,  LaVallie ER,  Collins-Racie LA,  Evans C;
PI   Merberg D,  Treacy M;
XX
DR   WPI; 2000-317937/27.
XX
PT   Isolated polynucleotides, and encoded proteins, comprising secreted
PT   expressed sequence tags (sESTs), useful for treating various disorders
PT   such as autoimmune, infectious, and central nervous system disorders -
XX
PS   Claim 1; Page 396; 618pp; English.
XX
CC   AAA41261 to AAA43419 represent specifically claimed secreted expressed
CC   sequence tags (sESTs), isolated from human, mouse, xenopus and rat
CC   tissue sources. The sESTs can have a range of activities depending on
CC   the tissues they were isolated from. The activities include:
CC   chemotactic; proliferative; immunomodulatory; haematopoietic;
CC   chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC   cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC   antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;
CC   nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC   anticonvulsant; and antidepressant. The sESTs can be used for gene
CC   therapy and in vaccines. The sESTs are useful as probes for the
CC   identification and isolation of full-length cDNAs and genomic DNA
CC   molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC   are useful in assays for determining biological activity and raising
CC   antibodies. They may be useful for treatment of autoimmune disorders
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression and
CC   psoriasis. AAA43420 to AAA43425 represent linker variants which are given
CC   in the exemplification of the present invention.
XX
SQ   Sequence 259 BP; 77 A; 49 C; 60 G; 68 T; 5 other;


  Query Match              2.3%;  Score 53.4;  DB 21;  Length 259;
  Best Local Similarity    53.3%;  Pred. No. 0.024;
  Matches  130;  Conservative    0;  Mismatches  111;  Indels    3;  Gaps    1;

Qy    1225 tcatataaagggcttcggagtgtttgtgtatatggtggtggaaatagagatgaacaaata 1284
           ||   ||  ||  | |  |||| |    |    |||||     ||  |      |||
Db      16 tcgactattggagtgcnaagtgctgtgattgtnggtggaattgattcaatntctcaatct 75

Qy    1285 gaagagcttaaaaaaggtgtagatatcataattgcaactcccggaagattgaatgatc-- 1342
             |  |||  ||||     | |||| ||||| |||||||| ||  || ||| ||| |
Db      76 ttggcccttgcaaaaaaaccacatataataatagcaactcctggtcgactgattgaccac 135

Qy    1343 -tgcaaatgagtaacttcgtcaatctgaagaatataacctacttggttttagatgaagca 1401
            || |||  |  ||     ||||  |||    | | |  |||||||  | |||||||||
Db     136 ttggaaaatacgaaaggtttcaacttgagagctctcaaatacttggtcatggatgaagcc 195

Qy    1402 gacaagatgttggacatgggatttgaaccccagataatgaagattttgttagatgtgcgc 1461
           |||   ||  || | | ||  |||||  |  || |    |||||  |   ||   | |
Db     196 gaccgaatantgaatanggattttgagacagaggttgacaagatcctcaaagtgattcct 255

Qy    1462 ccag 1465
           | ||
Db     256 cgag 259


RESULT   9
AAX84188
ID   AAX84188 standard; cDNA; 299 BP.
XX
AC   AAX84188;
XX
DT   08-SEP-1999  (first entry)
XX
DE   DNA encoding human breast tumour protein immunogenic fragment.
XX
KW   Breast tumour protein; immunogenic fragment; vaccine; detection;
KW   breast cancer development; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9933869-A2.
XX
PD   08-JUL-1999.
XX
PF   22-DEC-1998;   98WO-US27416.
XX
PR   17-JUL-1998;   98US-0118627.
PR   24-DEC-1997;   97US-0998253.
PR   24-DEC-1997;   97US-0998255.
PR   17-JUL-1998;   98US-0118554.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J;
XX
DR   WPI; 1999-405486/34.
```

```
XX
PT   New breast tumour protein genes used, in vaccines for immunotherapy,
PT   or for diagnosis of breast cancer
XX
PS   Claim 12; Page 54; 70pp; English.
XX
CC   This sequence encodes a human breast tumour protein immunogenic fragment
CC   of the invention. The polypeptides or nucleic acids encoding them are
CC   useful in vaccines and pharmaceutical compositions for manufacture of
CC   medicaments for inhibiting the development of breast cancer in a patient.
CC   They can also be used to treat breast cancer. Antibodies against these
CC   polypeptides can be used to detect and monitor progression of breast
CC   cancer in patients. Primers and probes derived from the polynucleotides
CC   encoding the breast proteins are useful for detection of breast cancer.
CC   Peripheral blood cells from a patient incubated in the presence of at
CC   least one polypeptide, such that T cells proliferate, are useful in
CC   manufacture of a medicament for treating breast cancer in a patient.
CC   Antigen presenting cells incubated in the presence of at least one
CC   polypeptide are also useful for treating breast cancer.
XX
SQ   Sequence 299 BP; 80 A; 69 C; 68 G; 79 T; 3 other;

  Query Match             2.3%;  Score 53.4;  DB 20;  Length 299;
  Best Local Similarity   54.8%;  Pred. No. 0.025;
  Matches  125;  Conservative    0;  Mismatches  102;  Indels    1;  Gaps    1;

Qy   1825 atactaattgcaactgatctagcctctagaggacttgatgtccatgacgttacacatgtc 1884
           || ||  | ||||||||  | |||    ||||  | ||  | | | | || |     ||
Db     31 attcttctagcaactgacgttgccagccgaggtttggacatacctcatgtaaatgtggtt 90

Qy   1885 tataattttgactttccacggaatattgaagaatacgtacaccgaatagggcgcacggga 1944
             || |||||| ||||     ||    | || ||| | || ||| |||| || || |
Db     91 gtcaactttgacattcctacccattccaaggattacatccatcgagtaggtcgaacagct 150

Qy   1945 agagcagggaggactggtgtttccattacaactttgactagaaatgat-tggagggttgc 2003
           ||||| ||| |  | ||     | |||||   | | ||     ||||| |||| | |
Db    151 agagctgggcgctccggaaaggctattacttttgtcacacagtatgatgtggaactcttc 210

Qy   2004 ctctgaattgattaatattctggaaagagcaaatcagagtattccaga 2051
           |   | || ||  | |    ||| ||||    | |||  | | | | |
Db    211 cagcgcatagaacacttnattgggaagaaactaccaggttttccaaca 258


RESULT  10
AAC79417
ID   AAC79417 standard; cDNA; 299 BP.
XX
AC   AAC79417;
XX
DT   07-FEB-2001  (first entry)
XX
DE   5' cDNA sequence of human breast tumour clone IS-5115.
XX
KW   Human; breast tumour antigen; cytostatic; immunotherapy;
KW   breast cancer; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061756-A2.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US09688.
XX
PR   09-APR-1999;   99US-0288950.
PR   02-JUL-1999;   99US-0346327.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC;
XX
DR   WPI; 2000-638568/61.
XX
PT   A novel isolated polypeptide comprising an immunogenic portion of a
PT   breast cancer protein useful in the detection and treatment of breast
PT   cancer -
XX
PS   Claim 13; Page 70; 95pp; English.
XX
CC   The present sequence was isolated from a breast tumour cDNA library. It
CC   is provided in a specification relating to compounds for immunotherapy
CC   and diagnosis of breast cancer. Breast tumour antigens and the
CC   polynucleotides that encode them may be used in the production of a
CC   pharmaceutical composition to be used in the treatment of breast cancer.
CC   Proliferated T cells and incubated antigen presenting cells are also
CC   required. The polypeptides and polynucleotides may also be used to
CC   produce a vaccine.
XX
SQ   Sequence 299 BP; 80 A; 69 C; 68 G; 79 T; 3 other;

  Query Match             2.3%;  Score 53.4;  DB 21;  Length 299;
  Best Local Similarity   54.8%;  Pred. No. 0.025;
  Matches  125;  Conservative    0;  Mismatches  102;  Indels    1;  Gaps    1;

Qy   1825 atactaattgcaactgatctagcctctagaggacttgatgtccatgacgttacacatgtc 1884
           || ||  | ||||||||  | |||    ||||  | ||  | | | | || |     ||
Db     31 attcttctagcaactgacgttgccagccgaggtttggacatacctcatgtaaatgtggtt 90

Qy   1885 tataattttgactttccacggaatattgaagaatacgtacaccgaatagggcgcacggga 1944
             || |||||| ||||     ||    | || ||| | || ||| |||| || || |
Db     91 gtcaactttgacattcctacccattccaaggattacatccatcgagtaggtcgaacagct 150

Qy   1945 agagcagggaggactggtgtttccattacaactttgactagaaatgat-tggagggttgc 2003
           ||||| ||| |  | ||     | |||||   | | ||     ||||| |||| | |
Db    151 agagctgggcgctccggaaaggctattacttttgtcacacagtatgatgtggaactcttc 210

Qy   2004 ctctgaattgattaatattctggaaagagcaaatcagagtattccaga 2051
           |   | || ||  | |    ||| ||||    | |||  | | | | |
Db    211 cagcgcatagaacacttnattgggaagaaactaccaggttttccaaca 258


RESULT  11
AAQ83249
ID   AAQ83249 standard; DNA; 235 BP.
XX
AC   AAQ83249;
XX
DT   15-FEB-1996  (first entry)
XX
DE   Streptococcus pneumoniae strain SPRU17 Exp9b DNA.
XX
KW   Exp9b; export protein; permease like protein;
KW   penicillin binding protein 1A; pbp1a; plpA; exp1; exp2; DEAD;
KW   pyruvate oxidase; virulence determinant; regulatory element;
KW   acellular vaccine; antibody; ds.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9506732-A2.
XX
PD   09-MAR-1995.
XX
PF   01-SEP-1994;   94WO-US09942.
XX
PR   01-SEP-1993;   93US-0116541.
PR   18-MAY-1994;   94US-0245511.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Masure HR,  Pearce BJ,  Tuomanen E;
XX
DR   WPI; 1995-115448/15.
DR   P-PSDB; AAR70161.
XX
```

PT Novel gene fragments encoding specific bacterial exported proteins
PT - specifically of S. pneumoniae, useful as vaccines
XX
PS Claim 19; Page 100-1; 168pp; English.
XX
CC This sequence represents the DNA encoding exp9b. The DNA encoding exp9a
CC can be seen in AAQ83248. Exp9 is a member of the D-E-A-D protein family
CC of ATP dependent RNA helicases. It is most similar to the DEAD from
CC E.coli. The protein sequences contain a conserved DEAD sequence within
CC the B motif of an ATP binding domain. Exp9 is an exported protein of
CC S.pneumoniae. Export proteins are the proteins in pathogenic bacteria
CC that are virulence determinants. Other export proteins include plpA
CC (see AAR70152), exp1, exp2 (see AAR70153), and pad1 (encoded by the
CC sequence shown in AAQ83259). This sequence can be inserted into an
CC expression vector (preferrably a bacterial expression vector) to provide
CC for high levels of expression of the protein. The protein can then be
CC used in the production of an acellular vaccine. These vaccines are used
CC to provide protection from Gram positive bacterial infection. Antibodies
CC against export proteins can be used for diagnosis of infection and in
CC passive immune therapy.
XX
SQ Sequence 235 BP; 58 A; 48 C; 58 G; 71 T; 0 other;

Query Match 2.1%; Score 50.8; DB 16; Length 235;
Best Local Similarity 51.3%; Pred. No. 0.087;
Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy   1679 ttgtcttcgtttctcgaaaagctgttgcggatcacttatcaagtgacctaatacttggaa 1738
          |||| || | |  |  |||    || |||| | ||  |  |||   | | | || |
Db      5 ttgtatttggtcgtaccaaacgccgtgtggatgaattgactcgtggtttgaaaattcgtg 64

Qy   1739 atatatcagtagagtctctgcatggagatagagaacagagagatcgggagaaagcattag 1798
               |    | |||     | ||||| ||   ||| || |   | ||       |   |
Db     65 gcttccgtgcagaaggaattcatggcgacctagaccaaaacaaacgtcttcgtgtccttc 124

Qy   1799 agaactttaaaacaggcaaagtgagaatactaattgcaactgatctagcctctagaggac 1858
               |||||||||  |||||  |     |  |  |||| || ||  | ||  |  | ||
Db    125 gtgactttaaaaatggcaatcttgatgttttggttgcgacagacgttgcagcgcgtggtt 184

Qy   1859 ttgatgtccatgacgttacacatgtctataattttgactttccacggaat 1908
          | ||| |    |  || || |||||||| || |  ||  ||||||   ||
Db    185 tggatatttcaggtgtgacccatgtctacaactacgatattccacaagat 234
```

RESULT 12
AAQ57421
ID AAQ57421 standard; cDNA to mRNA; 204 BP.
XX
AC AAQ57421;
XX
DT 19-OCT-1994 (first entry)
XX
DE DEAD ATP helicase like protein.
XX
KW Human cDNA; library; enzyme; protein; ss.
XX
OS Homo sapiens.
XX
PN WO9403599-A.
XX
PD 17-FEB-1994.
XX
PF 04-AUG-1993; 93WO-JP01095.
XX
PR 04-AUG-1992; 92JP-0208077.
PR 13-NOV-1992; 92JP-0327619.
PR 26-FEB-1993; 93JP-0061431.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;
XX
DR WPI; 1994-065688/08.
DR P-PSDB; AAR46082.
XX
PT cDNA of human origin and proteins coded by it - which may be
PT expressed by in vivo or in vitro translation using sense RNA or
PT antisense DNA corresponding to the cDNA.
XX
PS Claim 1; Page 37; 167pp; Japanese.
XX
CC mRNA expressed in human fibrosarcoma cell line HT-1080 was
CC isolated and used to construct a cDNA library using vector
CC pKA1. Clone HP00038 encoding DEAD ATP helicase-like protein
CC was isolated.
XX
SQ Sequence 204 BP; 51 A; 60 C; 54 G; 39 T; 0 other;

Query Match 2.0%; Score 47.2; DB 15; Length 204;
Best Local Similarity 62.9%; Pred. No. 0.51;
Matches 73; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy   1858 cttgatgtccatgacgttacacatgtctataattttgactttccacggaatattgaagaa 1917
          || ||  ||| || | |    || |||   |||| |||| | |||  | | |||||  |
Db      1 ctggacttccctgccatccaccacgtcatcaattatgacatgccagaggagattgagaac 60

Qy   1918 tacgtacaccgaatagggcgcacgggaagagcagggaggactggtgtttccattac 1973
          || |||||||| || || ||||| ||  |  | || |  || ||  |  ||| |||
Db     61 tatgtacaccggattggcgcaccgggcgctcgggaaacacaggcatcgccactac 116
```

RESULT 13
ABK09532
ID ABK09532 standard; cDNA; 102 BP.
XX
AC ABK09532;
XX
DT 14-MAR-2002 (first entry)
XX
DE Human ovarian tumour protein encoding cDNA #69.
XX
KW Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX
OS Homo sapiens.
XX
PN WO200190154-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US16895.
XX
PR 24-MAY-2000; 2000US-207107P.
PR 13-JUN-2000; 2000US-211457P.
PR 21-JUN-2000; 2000US-213673P.
PR 03-AUG-2000; 2000US-223288P.
PR 01-MAR-2001; 2001US-272790P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;
PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;
XX
DR WPI; 2002-097641/13.
XX
PT New isolated polynucleotide encoding polypeptide comprising portion of
PT ovarian tumour protein, useful for detection, diagnosis and therapy of
PT human ovarian cancer -
XX
PS Claim 1; Page 152; 285pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide comprising a portion of an ovarian tumour protein. The
CC sequences of the invention are useful for stimulating an immune response

```
CC   and for treating ovarian cancer in a patient. An antigen presenting cell
CC   that expresses the sequences is useful for treating ovarian cancer by
CC   incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells
CC   can then be proliferated and administered to the patient to inhibit the
CC   development of cancer. The DNA sequences are useful as probes or primers
CC   for nucleic acid hybridisation, to direct expression of a polypeptide in
CC   appropriate host cells. Detecting the presence of a cancer in a patient
CC   involves obtaining a biological sample from the patient, contacting the
CC   biological sample with an agent that binds to the protein, detecting the
CC   amount of protein that binds to the agent, comparing the amount of
CC   protein to a predetermined cut-off value and determining the presence of
CC   cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA
CC   molecules encoding ovarian tumour proteins of the invention.
XX
SQ   Sequence 102 BP; 24 A; 29 C; 28 G; 20 T; 1 other;


  Query Match             2.0%;  Score 47;  DB 24;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 0.47;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ttggtaccgagctcggatcctagtaacggccgccagtgtgctggaa 47
           ||||||||||||||||||||||||||||||||||||||||||||||
Db       9 ttggtaccgagctcggatcctagtaacggccgccagtgtgctggaa 55


RESULT  14
AAS49363/c
ID   AAS49363 standard; DNA; 197 BP.
XX
AC   AAS49363;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #587.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 1940; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 197 BP; 53 A; 40 C; 34 G; 70 T; 0 other;


  Query Match             2.0%;  Score 46.8;  DB 23;  Length 197;
  Best Local Similarity   57.5%;  Pred. No. 0.62;
  Matches   84;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy    1279 caaatagaagagcttaaaaaaggtgtagatatcataattgcaactcccggaagattgaat 1338
           |||||  |||   | ||||||||   | | ||| || | | ||| || ||  |  | |
Db     196 CAAATTAAAGCCTTGAAAAAAGGCCCACAAATCGTAGTCGGAACACCTGGGCGTGTTATC 137

Qy    1339 gatctgcaaatgagtaacttcgtcaatctgaagaatataacctacttggttttagatgaa 1398
           || |   ||  ||  |   | ||   | |    ||       ||| ||||||||||
Db     136 GACCATTTAAATCGTCGCACATTAAAAACGGACGGAATTCATACTTTGATTTTAGATGAA 77

Qy    1399 gcagacaagatgttggacatgggatt 1424
           || ||  | ||| || | ||||||||
Db      76 GCTGATGAAATGATGAATATGGGATT 51


RESULT  15
AAH71353/c
ID   AAH71353 standard; cDNA; 247 BP.
XX
AC   AAH71353;
XX
DT   19-SEP-2001 (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2627.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33312.
XX
PR   08-DEC-1999;   99US-0169681.
PR   21-DEC-1999;   99US-0171350.
PR   14-MAR-2000; 2000US-0189315.
PR   12-MAY-2000; 2000US-0203791.
PR   09-JUN-2000; 2000US-0210600.
PR   21-JUL-2000; 2000US-0220114.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer
PT   and for assessing and detecting compounds for treating the cancer -
XX
PS   Claim 1; Page 544; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
```

```
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy.
XX
SQ   Sequence 247 BP; 68 A; 31 C; 22 G; 126 T; 0 other;


  Query Match             2.0%;  Score 46.8;  DB 22;  Length 247;
  Best Local Similarity   52.6%;  Pred. No. 0.66;
  Matches  102;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy   2172 gggtagagaattcaagatttttttagaaatatagtaagacagaagtattggacatgttggc 2231
           ||| || || |  |   |||| |      | |||       ||   ||   |   ||
Db    226 GGGCAGGGATTGTATTTTTTTATTTTTTAAAAGTTTTTTTTAAAATTTTAGCTCCTTAAG 167

Qy   2232 agtatgaagagaccggactgatttgactgattcttaaaataatagtgtttgaaaatatag 2291
            | |  || ||       |   ||   ||  | |  ||| || ||| |||| || | |
Db    166 GGCAGAAAAAGGAACCTTTTTATTTTTTGCATTTCCAAAGAAAAGTTTTTGGCAAAAAAA 107

Qy   2292 aatccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaa 2351
           | |  |     |  |  ||   |||| ||||| | ||  |   |||   ||||||||||
Db    106 AGTTTAAAAAATGTTTTTTGAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 47

Qy   2352 aaaaaaaaaaaaa 2365
           |||||||||||||
Db     46 AAAAAAAAAAAAA 33


Search completed: September 24, 2002, 23:13:25
Job time: 11787 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:15:13 ; Search time 2465.76 Seconds
(without alignments)
12945.420 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 5562878

Minimum DB seq length: 0
Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	243.8	10.3	247	9	BE144935	BE144935 CM2-HT018
c	2	210.4	8.9	231	10	BF377169	BF377169 PM2-TN010
c	3	154.6	6.5	289	9	AA933999	AA933999 on91f04.s
	4	136.4	5.8	280	10	BJ066058	BJ066058 BJ066058
	5	115.6	4.9	289	9	BB068292	BB068292 BB068292
	6	103.8	4.4	270	10	BF153942	BF153942 022D01 Ma
	7	96.8	4.1	255	9	AA319350	AA319350 EST21563
	8	95.2	4.0	259	9	AA337335	AA337335 EST42053
	9	95	4.0	257	9	AA301156	AA301156 EST14135
	10	94	4.0	293	10	BI507376	BI507376 BB170007B
	11	92.8	3.9	290	9	AA370389	AA370389 EST82007
	12	92.4	3.9	248	9	AV256222	AV256222 AV256222
	13	91.2	3.9	254	10	BG688809	BG688809 336579 BA
c	14	76.6	3.2	287	9	AW390434	AW390434 RC0-ST018
	15	74	3.1	200	9	AA094227	AA094227 cl2141.se
	16	74	3.1	244	10	N31074	N31074 yx63e03.r1
	17	71.6	3.0	276	9	AW114510	AW114510 rs52b08.y
	18	70.2	3.0	300	9	BB161197	BB161197 BB161197
	19	69.2	2.9	223	10	BF598412	BF598412 sv17c05.y
c	20	68.4	2.9	277	9	AI641123	AI641123 fc19e02.x
	21	66.2	2.8	211	10	H49812	H49812 yq21e03.r1
	22	65.4	2.8	182	10	N44240	N44240 yy31f03.r1
	23	65.4	2.8	289	9	AA944137	AA944137 EST199636
	24	65.2	2.8	263	10	BG635374	BG635374 AT31944.5
	25	65	2.7	294	9	AV408925	AV408925 AV408925
	26	64.8	2.7	243	9	AA750925	AA750925 ISKH0250
	27	64.8	2.7	278	9	BB120001	BB120001 BB120001
	28	64.6	2.7	283	10	BF609454	BF609454 NXSI_044_
c	29	64.2	2.7	261	9	AW491036	AW491036 UI-M-BH3-
	30	63.4	2.7	249	10	BE528509	BE528509 M79D07STM
	31	63.2	2.7	225	10	Z17903	Z17903 ATTS0387 AC
	32	63.2	2.7	256	9	BE029873	BE029873 kp36a07.y
	33	62.8	2.7	205	9	AV348236	AV348236 AV348236
	34	62.2	2.6	283	9	AU173638	AU173638 AU173638
	35	62	2.6	219	10	W34581	W34581 mc58e03.r1
c	36	61.6	2.6	150	9	AA889106	AA889106 am38h05.s
	37	61.6	2.6	188	10	BM110125	BM110125 EST557661
c	38	60.8	2.6	234	9	AV589058	AV589058 AV589058
	39	60.6	2.6	200	10	BF376705	BF376705 RC6-TN007
	40	60.6	2.6	256	10	N65778	N65778 20818 Lambd
	41	60.6	2.6	280	9	BB597009	BB597009 BB597009
	42	60.4	2.6	218	9	AA358844	AA358844 EST67794
	43	60.2	2.5	293	10	BF010681	BF010681 NXCI_066_
	44	60	2.5	251	9	AA363887	AA363887 EST74365
	45	59.6	2.5	193	9	AA373151	AA373151 EST85145

ALIGNMENTS

RESULT 1
BE144935
LOCUS BE144935 247 bp mRNA linear EST 21-JUN-2000
DEFINITION CM2-HT0184-061099-018-g11 HT0184 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE144935
VERSION BE144935.1 GI:8607659
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 247)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM2-HT0184-061099-018-g11&t3=1999-10-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 247.
FEATURES Location/Qualifiers
source 1. .247
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone_lib="HT0184"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 109 a 39 c 48 g 51 t
ORIGIN

Query Match 10.3%; Score 243.8; DB 9; Length 247;
Best Local Similarity 99.2%; Pred. No. 4.1e-27;
Matches 245; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    462 tggtgggtcaaaaataaagaatatacaaagtacaacaaacaccacaatccaaataataca 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGGTGGGTCAAAAATAAAGAATATACAAAGTACAACAAACACCACAATCCAAATAATACA 60

Qy    522 agaacaaccagaatcattagtcaaaatttttggcagcaaggcaatgcaaacgaaagcaaa 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGAACAACCAGAATCATTAGTCAAAATTTTTGGCAGCAAGGCAATGCAAACGAAAGCAAA 120

Qy    582 agcagtgatagacaattttgttaaaaagctagaagaaaattacaattcagaatgcggaat 641
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGCAGTGATAGACAATTTGGTTAAAAAGCTAGAAGAAAATTACAATTCAGAATGCGGAAT 180

Qy    642 tgatactgcattccaaccttctgttggaaaagatggaagcacagataacaatgttgttgc 701
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    181 TGATACTGCATTCCAACCTTCTGTGGGAAAAGATGGAAGCACAGATAACAATGTTGTTGC 240

Qy    702 aggagat 708
          |||||||
Db    241 AGGAGAT 247
```

RESULT 2
BF377169/c
LOCUS BF377169 231 bp mRNA linear EST 24-NOV-2000
DEFINITION PM2-TN0103-040900-001-c02 TN0103 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF377169
VERSION BF377169.1 GI:11339194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 231)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM2&t2=PM2-TN0103-040900-001-c02&t3=2000-09-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 231.
FEATURES Location/Qualifiers
source 1. .231
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TN0103"
/dev_stage="Adult"
/note="Organ: testis_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 52 a 46 c 38 g 95 t
ORIGIN

Query Match 8.9%; Score 210.4; DB 10; Length 231;
Best Local Similarity 97.0%; Pred. No. 4.1e-22;
Matches 225; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy    497 caaacaccacaatccaaataatacaagaacaaccagaatcattagtcaaaatttttggca 556
          |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    231 CAAACACCGCAATCTAAATAATACAAGAACAACCAGAATCATTAGTCAAAATTTTTGGCA 172

Qy    557 gcaaggcaatgcaaacgaaagcaaaagcagtgatagacaattttgttaaaaagctagaag 616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 GCAAGGCAATGCAAACGAAAGCAAAAGCAGTGATAGACAATTTTGTTAAAAAGCTAGAAG 112

Qy    617 aaaattacaattcagaatgcggaattgatactgcattccaaccttctgttggaaaagatg 676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 AAAATTACAATTCAGAATGCGGAATTGATACTGCATTCCAACCTTCTGTTGGAAAAGATG 52

Qy    677 gaagcacagataacaatgttgttgcaggagatcggccattgatagattggga 728
          ||||| |||||||||||||||||||||||||  ||  ||||||||||||
Db     51 GAAGCTCAGATAACAATGTTGTTGCAGGAGATCGTACA-CGATAGATTGGGA 1
```

RESULT 3
AA933999/c
LOCUS AA933999 289 bp mRNA linear EST 23-JUN-1998
DEFINITION on91f04.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1564063 3' similar to TR:O13370 O13370 SUPPRESSOR OF UNCONTROLLED MITOSIS. ;, mRNA sequence.
ACCESSION AA933999
VERSION AA933999.1 GI:3090267
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 289)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 608 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1564063"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 83 a 67 c 51 g 88 t
ORIGIN

Query Match 6.5%; Score 154.6; DB 9; Length 289;
Best Local Similarity 71.9%; Pred. No. 8.3e-14;
Matches 202; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy   1793 cattagagaactttaaaacaggcaaagtgagaatactaattgcaactgatctagcctcta 1852
          |||||||||  ||||||   || ||  | |  ||||| ||| ||||||||| | |  ||
Db    288 CATTAGAGACTTTTAAAGGCGGAAACTTAAAGATACTGATTACAACTGATTTGGTTTCCC 229

Qy   1853 gaggacttgatgtccatgacgttacacatgtctataattttgactttccacggaatattg 1912
          ||||  |||||| |  |||| || |||||||| | |||||  ||| || |||  ||||| |||
Db    228 GAGGTCTTGATCTTAATGATGTCACACATGTTTTTAATTATGATTTCCCAAGGAATTTTG 169

Qy   1913 aagaatacgtacaccgaatagggcgcacgggaagagcagggaggactggtgtttccatta 1972
          | | |||  |||||| || |||||   ||  |||  |  | || | ||||||    ||  ||
Db    168 ACGTATATGTACACAGAGTAGGGTACATTGGACGGACGGGAAAGACTGGCACATCATTTC 109

Qy   1973 caactttgactagaaatgattggagggttgcctctgaattgattaatattctggaaagag 2032
          |     | |||  |  ||||  ||   | |||  |||||||||||| ||||||||  ||||
Db    108 CCCTCATCACTCGGAGGGATTCGAAAATGGCCGGTGAATTGATTAAAATTCTGGACAGAG 49

Qy   2033 caaatcagagtattccagaggagcttgtatcaatggctgag 2073
          ||||||||||  ||||  ||  ||  ||||||   |||||||||
Db     48 CAAATCAGAGTGTTCCGGAAGATCTTGTAGTAATGGCTGAG 8
```

RESULT 4
BJ066058
LOCUS BJ066058 280 bp mRNA linear EST 10-DEC-2001
DEFINITION BJ066058 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL085a10 5', mRNA sequence.
ACCESSION BJ066058
VERSION BJ066058.1 GI:17479113
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 280)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .280
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL085a10"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 81 a 59 c 66 g 73 t 1 others
ORIGIN

Query Match 5.8%; Score 136.4; DB 10; Length 280;
Best Local Similarity 71.5%; Pred. No. 4.4e-11;
Matches 193; Conservative 0; Mismatches 76; Indels 1; Gaps 1;

```
Qy    982 aaggcaggttttcaaaagccaacacctattcagt-cacaggcatggcccattgtgttgca 1040
          || || || ||| | | ||| || || || |||| |||||||||||||| ||| |  ||||
Db      9 AAAGCTGGATTTGAGAGGCCCACCCCAATACAGTNCACAGGCATGGCCAATTATACTGCA 68

Qy   1041 aggaatagatcttataggagtagcccagactggaacaggaaagacattgtgttatttaat 1100
          ||| | |||||||| ||| | ||||||||||| || ||||| ||||||    ||||| |
Db     69 GGGATTTGATCTTATTGGAATTGCCCAGACTGGCACTGGAAAAACATTGGCACATTTACT 128

Qy   1101 gcctggatttattcatctggtccttcaacccagccttaaaggtcaaaggaatagacccgg 1160
          || ||||| |||||| |||   | ||||| |  |  | ||  ||| |  || | || ||
Db    129 TCCAGGATTCATTCATTTGGATTTGCAACCAATTCCCAGAGAGCAACGAGATGGGCCTGG 188

Qy   1161 catgttagttctaactcccactcgggaattagcacttcaagtagaaggagaatgttgcaa 1220
          |||||| |||||  |||| ||||| ||||| || |||||||  | |  || || |  ||
Db    189 CATGTTGGTTCTTGCTCCAACTCGTGAATTGGCCCTTCAAGTTAAGGCTGAGTGCTCAAA 248

Qy   1221 atattcatataaagggcttcggagtgtttg 1250
          ||   |||||||||  ||   ||| ||||
Db    249 GTACAAATATAAAGGATTTGAAAGTATTTG 278
```

RESULT 5
BB068292
LOCUS BB068292 289 bp mRNA linear EST 25-JUN-2000
DEFINITION BB068292 RIKEN full-length enriched, 15 days embryo male testis Mus musculus cDNA clone 8030471J02 3', mRNA sequence.
ACCESSION BB068292
VERSION BB068292.1 GI:8525706
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 289)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki ,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="8030471J02"
                     /clone_lib="RIKEN full-length enriched, 15 days embryo
                     male testis"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="15 days embryo"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of subtraction to
                     Rot = 185.0 Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT      105 a     44 c     60 g     80 t
ORIGIN
```

```
  Query Match             4.9%;  Score 115.6;  DB 9;  Length 289;
  Best Local Similarity   67.7%;  Pred. No. 5.6e-08;
  Matches  199;  Conservative    0;  Mismatches   79;  Indels   16;  Gaps    2;

Qy   2029 agagcaaatcagagtattccagaggagcttgtatcaatggctgagaggtttgaggcacat 2088
          ||| || |||||  | | |||| ||  ||| | | |||| || |||| |   | ||| ||
Db      1 AGACCACATCAGCATTTCCCAGTGGTACTTCTCTTAATGCCTAAGAGATACAAAGCAAAT 60

Qy   2089 caacggaaaagggaaatggaaagaaaaatggaaagacctcaaggaaggcccaagaagttt 2148
           |||   |||| || ||||||| |||||||| |||||||||||||| |||| ||||||||
Db     61 AAACTACAAAGAGAGATGGAAAAAAAAATGGGAAGACCTCAAGGAAAGCCCCAGAAGTTT 120

Qy   2149 cattaatgtcttctgtactagtggggtagagaattcaagatttttagaaatatagtaag 2208
           | ||||||    || |  | | ||      |       |  || |||  | |  |
Db    121 TACTAATGTTGCGTGGAGCACTAGGTGCCGCAGGATTCAAGATTGTAGTCAAAGTG---- 176

Qy   2209 acagaagtattggacatgttggcagtatgaagagaccggactgatttgactgattcttaa 2268
             ||||| ||| |      | || |||| ||||||          ||| ||||||||||
Db    177 ---GAAGTGTTGAAACATACGTCACTATGGAGAGAC---------TTGGCTGATTCTTAA 224

Qy   2269 aataatagtgtttgaaaatatagaatccagtgttttatactttctttaataaaa 2322
          |||||||||||| | ||||| |||||| || ||||||||||||||||||||||||
Db    225 AATAATAGTGTGTTAAAATGTAGAATTCAATGTTTTATACTTTCTTTAATAAAA 278
```

```
RESULT   6
BF153942
LOCUS       BF153942                270 bp    mRNA    linear   EST 01-DEC-2000
DEFINITION  022D01 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to
            RNA HELICASE-LIKE PROTEIN DB10 gi|1084413|. . . sp|P46942, mRNA
            sequence.
ACCESSION   BF153942
VERSION     BF153942.1  GI:11035882
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   Nielsen,K.L, Crookshanks,M., Emmersen,J. and Welinder,K.G.
  TITLE     EST-sequencing of mature potato tuber (Var. Kuras)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Karen G. Welinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk
            High quality sequence stop: 270
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .270
                     /organism="Solanum tuberosum"
                     /cultivar="Field grown Kuras"
                     /db_xref="taxon:4113"
                     /clone_lib="Mature tuber lambda ZAP"
                     /tissue_type="Tuber"
                     /note="Vector: Lambda ZAP"
BASE COUNT       73 a     45 c     80 g     72 t
ORIGIN
```

```
  Query Match             4.4%;  Score 103.8;  DB 10;  Length 270;
  Best Local Similarity   63.3%;  Pred. No. 3.3e-06;
  Matches  159;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy   1249 tgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtagat 1308
          ||  | |||||| ||||     | || |   ||| |   |||||| |  |  |||||||||
Db     11 TGCCTGTATGGAGGTGCTCCAAAAGGTCCTCAACTCCGAGAGCTGAGTAGGGGTGTAGAT 70

Qy   1309 atcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatctg 1368
           ||  |  | || |||||| ||  | ||||||||  |       |  |    | | |||
Db     71 ATTGTTGTCGCCACTCCTGGTCGTTTGAATGACATATTGGAAATGAGAAGGCTTAGTCTT 130

Qy   1369 aagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttgaa 1428
                |  |  | ||  |||| ||||||||||||||    ||| | ||||||||||||||
Db    131 GGTCAAGTTTCTTAGCTGGTGTTAGATGAAGCAGATCGTATGCTAGACATGGGATTTGAA 190

Qy   1429 ccccagataatgaagattttgttagatgtgcgcccagataggcagacagttatgaccagt 1488
          || |||||||  |||||| ||  ||| ||||   |  | |||||||||  | |||   |
Db    191 CCTCAGATAAGAAAGATTGTGAAAGAGGTGCCTACGCAGAGGCAGACATTGATGTATACA 250

Qy   1489 gctacatggcc 1499
          | |||||||||
Db    251 GATACATGGCC 261
```

```
RESULT   7
AA319350
LOCUS       AA319350                255 bp    mRNA    linear   EST 19-APR-1997
DEFINITION  EST21563 Adrenal gland tumor Homo sapiens cDNA 5' end similar to
            similar to RNA helicase p68, mRNA sequence.
ACCESSION   AA319350
VERSION     AA319350.1  GI:1971676
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 255)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
```

,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W. , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 96026280

COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
source 1. .255
/organism="Homo sapiens"
/db_xref="ATCC (inhost):119896"
/db_xref="taxon:9606"
/clone_lib="Adrenal gland tumor"
/dev_stage="adult"
/note="Organ: adrenal gland; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 80 a 42 c 63 g 66 t 4 others

ORIGIN

Query Match 4.1%; Score 96.8; DB 9; Length 255;
Best Local Similarity 65.3%; Pred. No. 3.7e-05;
Matches 154; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

```
Qy   1248 ttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtaga 1307
          |||| | || |||||||    ||  |    |||||||    |  |  | | |||||| ||
Db     17 TTGTATCTACGGTGGTGCTCCTAAGGGACCACAAATACGTNATTTGGAGAGAGGTGTGGA 76

Qy   1308 tatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatct 1367
           |||   |||||||| || |||||| ||| |||| |   |  | ||        ||||||
Db     77 AATCTGTATTGCAACACCTGGAAGACTGATTGATTTTTNAGAGTGTGGAAAAACCAATCT 136

Qy   1368 gaagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttga 1427
           |  |  | |||||||| | ||  | ||||||||||| |  ||| | || ||||| |||||
Db    137 NAGAAGAACAACCTACCTTGTCCTTGATGAAGCAGATAGAATGCTTGANATGGGCTTTGA 196

Qy   1428 accccagataatgaagattttgtta-gatgtgcgcccagataggcagacagttatg 1482
          |||||| |||| ||||||| ||  |  |  |  | || |||||||| ||  | |||
Db    197 ACCCCAAATAAGGAAGATTGTGGGATCAAATAAGACCTGATAGGCAAACTCTAATG 252
```

RESULT 8
AA337335
LOCUS AA337335 259 bp mRNA linear EST 21-APR-1997
DEFINITION EST42053 Endometrial tumor Homo sapiens cDNA 5' end similar to similar to RNA helicase p68, mRNA sequence.
ACCESSION AA337335
VERSION AA337335.1 GI:1989573
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 259)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W. , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE 96026280

COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
source 1. .259
/organism="Homo sapiens"
/db_xref="ATCC (inhost):139075"
/db_xref="taxon:9606"
/clone_lib="Endometrial tumor"
/sex="female"
/dev_stage="adult"
/note="Organ: endometrium; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 80 a 42 c 66 g 68 t 3 others

ORIGIN

Query Match 4.0%; Score 95.2; DB 9; Length 259;
Best Local Similarity 65.1%; Pred. No. 6.5e-05;
Matches 153; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

```
Qy   1248 ttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtaga 1307
          |||| | || |||||||    ||  |    |||||||   ||  |  | | |||||| ||
Db     17 TTGTATCTACGGTGGTGCTCCTAAGGGACCACAAATACGTGATTTGGAGAGAGGTGTGGA 76

Qy   1308 tatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatct 1367
           |||   |||||||| || |||||| ||| |||  |   |  | ||        ||||||
Db     77 AATCTGTATTGCAACACCTGGAAGACTGATTGANTTTTAGAGTGTGGAAAAACCAATCT 136

Qy   1368 gaagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttga 1427
          ||  |    ||||||| | ||  | ||||||||||| |  ||| | || ||||| |||||
Db    137 GAGAAGANCAACCTACCTTGTCCTTGATGAAGCAGATAGAATGCTTGANATGGGCTTTGA 196

Qy   1428 accccagataatgaagatt-ttgttagatgtgcgcccagataggcagacagttat 1481
          |||||| |||| ||||||| | | |  |  |  | || |||||||| ||  | ||
Db    197 ACCCCAAATAAGGAAGATTGTGGATCAAAATAAGACCTGATAGGCAAACTCTAAT 251
```

RESULT 9
AA301156

```
LOCUS       AA301156                257 bp    mRNA    linear   EST 18-APR-1997
DEFINITION  EST14135 Testis tumor Homo sapiens cDNA 5' end similar to similar
            to RNA helicase p68, mRNA sequence.
ACCESSION   AA301156
VERSION     AA301156.1  GI:1953552
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 257)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .257
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):191788"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis tumor"
                     /sex="male"
                     /dev_stage="adult"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       79 a     42 c     64 g     68 t      4 others
ORIGIN


  Query Match             4.0%;  Score 95;  DB 9;  Length 257;
  Best Local Similarity   64.2%;  Pred. No. 6.9e-05;
  Matches  140;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy   1248 ttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtaga 1307
          |||| | || |||||||    ||  |    |||||||  ||  |  | | |||||| ||
Db     17 TTGTATCTACGGTGGTGCTCCTAAGGGACCACAAATACGTGATTTGGAGAGAGGTGTGGA 76

Qy   1308 tatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatct 1367
           |||   |||||||| || |||||| ||| |||  |   |  | ||        ||||||
Db     77 AATCTGTATTGCAACACCTGGAAGACTGATTGANTTTTTAGAGTGTGGAAAAACCAATCT 136

Qy   1368 gaagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttga 1427
          ||  |  | |||||||  | ||  | ||||||||||| |  ||| | || |||||  |||||
Db    137 GAGAAGAACAACCTACCTTGTCCTTGATGAAGCAGATAGAATGCTTGATATGGGCTTTGA 196
```

```
Qy   1428 accccagataatgaagattttgttagatgtgcgcccag 1465
          |||||| |||| |||| || ||    |  |  | || |
Db    197 ACCCCAAATAAGGAAGNTTGTGGATCAAATAAGNCCTG 234


RESULT  10
BI507376
LOCUS       BI507376                293 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  BB170007B10A05 Bee Brain Normalized/Subtracted Library, BB17 Apis
            mellifera cDNA clone BB170007B10A05 5', mRNA sequence.
ACCESSION   BI507376
VERSION     BI507376.1  GI:15357750
KEYWORDS    EST.
SOURCE      honeybee.
  ORGANISM  Apis mellifera
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata
            ; Apoidea; Apidae; Apis.
REFERENCE   1  (bases 1 to 293)
  AUTHORS   Whitfield,C.W., Soares,B., Robertson,H.M., Pardinas,J., Liu,L.,
            Smoller,D. and Robinson,G.E.
  TITLE     An Expressed Sequence Tag Resource for Studies of Brain and
            Behavior in the Honey Bee
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Gene E. Robinson
            Department of Entomology
            University of Illinois
            505 S. Goodwin Ave., Urbana, IL 61801, USA
            Tel: 217 265 0309
            Fax: 217 244 3499
            Email: generobi@life.uiuc.edu
            This research was funded by the University of Illinois Critical
            Research Initiatives Fund and a Burroughs-Wellcome Trust Innovation
            Award in Functional Genomics to G.E. Robinson and an NSF
            Postdoctoral Fellowship in Bioinformatics to C.W. Whitfield.
            PCR PRimers
            FORWARD: TAATACGACTCACTATAGGG
            BACKWARD: ATTAACCCTCACTAAAG
            Insert Length: 293   Std Error: 0.00
            Plate: BB170007B10  row: A  column: 05
            Seq primer: AGCGGATAACAATTTCACACAGGA
            High quality sequence stop: 293.
FEATURES             Location/Qualifiers
     source          1. .293
                     /organism="Apis mellifera"
                     /strain="mixed strains of European bees, predominantly
                     A.m. ligustica"
                     /db_xref="taxon:7460"
                     /clone="BB170007B10A05"
                     /clone_lib="Bee Brain Normalized/Subtracted Library, BB17"
                     /sex="female"
                     /tissue_type="brain"
                     /dev_stage="adult worker honey bee"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; This BB17 cDNA library was generated by
                     subtraction of the BB16  library with 4000 previously
                     sequenced clones. The BB16 library was  contributed by the
                     Soares laboratory and it was constructed and  normalized
                     as described by Bonaldo, M.F., Lennon, G. and  Soares,
                     M.B. (1996), Genome Research 6(9): 791-806. RNA was
                     prepared  from dissected brains of adult worker bees of
                     various ages and  various behavioral groups. "
BASE COUNT       97 a     38 c     69 g     89 t
ORIGIN


  Query Match             4.0%;  Score 94;  DB 10;  Length 293;
  Best Local Similarity   59.3%;  Pred. No. 9.3e-05;
  Matches  160;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy   1241 ggagtgtttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaag 1300
           | ||   |||| | | |||||||| |   | ||     |||      ||  |  || |||
```

```
Db     20 GAAGCACTTGTATTTTTGGTGGTGCACCAAAAGGAAGCCAAGCTCGTGATTTAGAACAAG 79

Qy   1301 gtgtagatatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcg 1360
          | || || ||    ||||| ||||| ||| |||| | |||| |         || |
Db     80 GAGTTGAGATTTGCATTGCTACTCCTGGACGATTAATTGATTTTTTGGAACGTGGCACAA 139

Qy   1361 tcaatctgaagaatataacctacttggttttagatgaagcagacaagatgttggacatgg 1420
            ||| |    |     || || ||||| || ||||||||||| |  ||||| || ||||
Db    140 CTAATTTACGTAGATGTACATATTTGGTATTGGATGAAGCAGATAGAATGTTAGATATGG 199

Qy   1421 gatttgaaccccagataatgaagattttgttagatgtgcgcccagataggcagacagtta 1480
          | |||||||| ||||| |  || ||| |   | |  |  | |||||||| ||   | | |
Db    200 GTTTTGAACCACAGATTAGAAAAATTATTGAACAAATCAGACCAGATAGACAAGTATTAA 259

Qy   1481 tgaccagtgctacatggcctcattcagttc 1510
          ||     ||| || |||||  |   |||||
Db    260 TGTGGTCTGCAACTTGGCCAAAAGAAGTTC 289


RESULT  11
AA370389
LOCUS       AA370389                 290 bp    mRNA    linear   EST 21-APR-1997
DEFINITION  EST82007 Prostate gland I Homo sapiens cDNA 5' end similar to
            similar to RNA helicase p68, mRNA sequence.
ACCESSION   AA370389
VERSION     AA370389.1  GI:2022706
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 290)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .290
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):174856"
                     /db_xref="taxon:9606"
                     /clone_lib="Prostate gland I"
                     /sex="male"
                     /dev_stage="adult, 21 yrs"
                     /note="Organ: prostate; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       89 a     52 c     75 g     74 t
ORIGIN


  Query Match             3.9%;  Score 92.8;  DB 9;  Length 290;
  Best Local Similarity   66.5%;  Pred. No. 0.00014;
  Matches  133;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy   1248 ttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtaga 1307
          |||| | || |||||||    ||  |    |||||||   ||  |  | | |||||| ||
Db     71 TTGTATCTACGGTGGTGCTCCTAAGGGACCACAAATACGTGATTTGGAGAGAGGTGTGGA 130

Qy   1308 tatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatct 1367
            |||   |||||||| || |||||| ||| |||  |   |  | ||         ||||||
Db    131 AATCTGTATTGCAACACCTGGAAGACTGATTGACTTTTTAGAGTGTGGAAAAACCAATCT 190

Qy   1368 gaagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttga 1427
          ||  |  | |||||||| | ||  | |||||||||||| |  ||| | || |||||  |||||
Db    191 GAGAAGAACAACCTACCTTGTCCTTGATGAAGCAGATAGAATGCTTGATATGGGCTTTGA 250

Qy   1428 accccagataatgaagattt 1447
          ||||||  ||||  ||||  |||
Db    251 ACCCCAAATAAGGAAGGTTT 270


RESULT  12
AV256222
LOCUS       AV256222                 248 bp    mRNA    linear   EST 04-NOV-1999
DEFINITION  AV256222 RIKEN full-length enriched, adult male testis (DH10B) Mus
            musculus cDNA clone 4921526D21 3', mRNA sequence.
ACCESSION   AV256222
VERSION     AV256222.1  GI:6243681
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 248)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
            Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
            Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
            ,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
            Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
            Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
            Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y.,
            Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
            Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al. 1999)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
            Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
            ,Y.
             Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
```

```
           Please visit our web site (http://genome.rtc.riken.go.jp) for
           further details.
FEATURES             Location/Qualifiers
     source          1. .248
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4921526D21"
                     /clone_lib="RIKEN full-length enriched, adult male testis
                     (DH10B)"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
BASE COUNT       90 a     33 c     53 g     72 t
ORIGIN

  Query Match             3.9%;  Score 92.4;  DB 9;  Length 248;
  Best Local Similarity   67.2%;  Pred. No. 0.00017;
  Matches  168;  Conservative    0;  Mismatches   66;  Indels   16;  Gaps    2;

Qy   2073 gaggtttgaggcacatcaacggaaaagggaaatggaaagaaaaatggaaagacctcaagg 2132
           ||| |   | ||| || ||| ||||| || |||||||  ||||||  ||||||||||||
Db      4 GAGATACAAAGCATATAAACTTAAAAGAGAGATGGAAAATCAAATGGGAAGACCTCAAGG 63

Qy   2133 aaggcccaagaagtttcattaatgtcttctgtactagtggggtagagaattcaagatttt 2192
           || |||| |||||||| | ||||||    || |  | | ||      |       |  ||
Db     64 AAAGCCCCAGAAGTTTTACTAATGTTGCGTGGAGCACTAGGTCCCGCAGGATTCAAGATT 123

Qy   2193 ttagaaatatagtaagacagaagtattggacatgttggcagtatgaagagaccggactga 2252
           |||  | |  |       ||||| ||| |      | || |||| ||||||
Db    124 GTAGTCAAAGTG-------GAAGTGTTGAAACATACGTCACTATGGAGAGAC-------- 168

Qy   2253 tttgactgattcttaaaataatagtgtttgaaaatatagaatccagtgttttatactttc 2312
           ||| ||||||||||||||||||||||| | ||||| |||||| || ||||||||||||||
Db    169 -TTGGCTGATTCTTAAAATAATAGTGTGTTAAAATGTAGAATTCAATGTTTTATACTTTC 227

Qy   2313 tttaataaaa 2322
           ||||||||||
Db    228 TTTAATAAAA 237


RESULT  13
BG688809
LOCUS       BG688809                 254 bp    mRNA    linear   EST 02-MAY-2001
DEFINITION  336579 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BG688809
VERSION     BG688809.1  GI:13930610
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 254)
  AUTHORS   Sonstegard,T.S., Capuco,A.V., Van Tassell,C.P., Ashwell,M.S. and
            Wells,K.D.
  TITLE     Mapping of Expressed Sequence Tags from a normalized bovine mammary
            gland cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Sonstegard TS
            USDA, ARS, Beltsville Agricultural Research Center
            Bdlg. 200 Rm 2A, Beltsville, MD 20705, USA
            Tel: 301 504 8416
            Fax: 301 504 8414
            Email: tads@lpsi.barc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 36  row: K  column: 10
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .254
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="BARC 5BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled mRNA isolated from mammary
                     tissues at eight physiological, developmental, and disease
                     states."
BASE COUNT       78 a     47 c     68 g     61 t
ORIGIN

  Query Match             3.9%;  Score 91.2;  DB 10;  Length 254;
  Best Local Similarity   62.7%;  Pred. No. 0.00026;
  Matches  158;  Conservative    0;  Mismatches   93;  Indels    1;  Gaps    1;

Qy   1248 ttgtgtatatggtggtggaaatagagatgaacaaatagaagagcttaaaaaaggtgtaga 1307
           ||| || ||||| ||||     |  |    |||||||   ||  |  ||| ||||||  ||
Db      1 TTGCGTTTATGGCGGTGCTCCCAAGGGACCACAAATACTTGATTTGGAAAGAGGTGTGGA 60

Qy   1308 tatcataattgcaactcccggaagattgaatgatctgcaaatgagtaacttcgtcaatct 1367
           |||   |||| ||| || |||||| |||  ||  |   |    ||        ||||||
Db     61 AATCTGTATTG-AACACCTGGAAGACTGATAGACTTTTTAGAATGTGGGAAAACCAATCT 119

Qy   1368 gaagaatataacctacttggttttagatgaagcagacaagatgttggacatgggatttga 1427
           |  |  |  |||||| | ||  | ||||||||||      || | || ||||| |||||
Db    120 AAGAAGAACCACCTACCTTGTCCTTGATGAAGCAGGGGGAGTGCTTGATATGGGGTTTGA 179

Qy   1428 accccagataatgaagattttgttagatgtgcgcccagataggcagacagttatgaccag 1487
           |||||| |||| ||||||| ||    |  |  | || || ||||| ||  | |||   ||
Db    180 ACCCCAAATAAGGAAGATTGTGGACCAGATCAGACCTGACAGGCAAACCCTAATGTGGAG 239

Qy   1488 tgctacatggcc 1499
           |||||| |||||
Db    240 TGCTACTTGGCC 251


RESULT  14
AW390434/c
LOCUS       AW390434                 287 bp    mRNA    linear   EST 04-FEB-2000
DEFINITION  RC0-ST0185-231099-032-f03 ST0185 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW390434
VERSION     AW390434.1  GI:6895093
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 287)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC0&t2=RC0-ST0185-
            231099-032-f03&t3=1999-10-23&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 47
            High quality sequence stop: 287.
FEATURES              Location/Qualifiers
     source           1. .287
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="ST0185"
                      /dev_stage="Adult"
                      /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
BASE COUNT       72 a     71 c     42 g    102 t
ORIGIN


  Query Match             3.2%;  Score 76.6;  DB 9;  Length 287;
  Best Local Similarity   63.1%;  Pred. No. 0.037;
  Matches  118;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy    1391 tagatgaagcagacaagatgttggacatgggatttgaaccccagataatgaagatttgt 1450
           | ||||||||||| |  ||| | || ||||| |||||||||||| |||| ||||||| ||
Db     285 TTGATGAAGCAGATAGAATGCTTGATATGGGCTTTGAACCCCAAATAAGGAAGATTGTGG 226

Qy    1451 tagatgtgcgcccagataggcagacagttatgaccagtgctacatggcctcattcagttc 1510
           |  |  | |  |||||||| ||  | |||   ||||| || |||||  |   |||
Db     225 ATCAAATAAGACGTGATAGGCAAACTCTAATGTGGAGTGCGACTTGGCCAAAAGAAGTAA 166

Qy    1511 atcgcctcgcacaatcttatttgaaagaaccaatgattgtctatgttggtacattggatc 1570
           |  || ||  ||  ||   |||||||    ||   | |  |  ||||| || | || |
Db     165 GACAGCTTGCTGAAGATTTCCTGAAAGACTATATTCATATAAACATTGGTGCACTTGAAC 106

Qy    1571 tagttgc 1577
           |   |||
Db     105 TGAGTGC 99


RESULT  15
AA094227
LOCUS       AA094227                 200 bp    mRNA    linear   EST 25-OCT-1996
DEFINITION  cl2141.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
            cDNA 5', mRNA sequence.
ACCESSION   AA094227
VERSION     AA094227.1  GI:1639804
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 200)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES              Location/Qualifiers
     source           1. .200
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="Human fetal heart, Lambda ZAP Express"
                      /lab_host="E. coli XL1-Blue"
                      /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                      XhoI; mRNA was purified from human fetal hearts (8-10
                      weeks).  cDNA was synthesized using a XhoI-Oligo dT
                      adaptor-primer.  EcoRI adaptors were ligated, followed by
                      digestion with XhoI, for directional cloning into
                      predigested lambda ZAP Express."
BASE COUNT       67 a     34 c     50 g     49 t
ORIGIN


  Query Match             3.1%;  Score 74;  DB 9;  Length 200;
  Best Local Similarity   64.1%;  Pred. No. 0.1;
  Matches  127;  Conservative    0;  Mismatches   70;  Indels    1;  Gaps    1;

Qy    1278 acaaatagaagagcttaaaaaaggtgtagatatcataattgcaactcccggaagattgaa 1337
           |||||||   ||  |  | | |||||| || |||   |||||||| || |||||| |||
Db       4 ACAAATACGTGATTTGGAGAGAGGTGTGGAAATCTGTATTGCAACACCTGGAAGACTGAT 63

Qy    1338 tgatctgcaaatgagtaacttcgtcaatctgaagaatataacctacttggttttagatga 1397
           |||  |   |  | ||       ||||||||  |  | |||||||  | ||  | ||||
Db      64 TGACTTTTTAGAGTGTGGAAAAACCAATCTGAGAAGAACAACCTACCTTGTCCTTGATG- 122

Qy    1398 agcagacaagatgttggacatgggatttgaaccccagataatgaagattttgttagatgt 1457
           ||||||  |  ||| | || ||||| |||||||||||| |||| | ||  ||    |   |
Db     123 AGCAGATAGAATGCTTGATATGGGCTTTGAACCCCAAATAAGGGAGGATTGTGGATCAAT 182

Qy    1458 gcgcccagataggcagac 1475
           |  ||  |||||||  ||
Db     183 AAGACCTCATAGGCAAAC 200


Search completed: September 24, 2002, 21:55:44
Job time: 9631 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:12:38 ; Search time 4252.24 Seconds
(without alignments)
11638.868 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 843946

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	39.4	1.7	99	6	A08898	A08898 H.sapiens (
	2	39.2	1.7	86	6	AX203294	AX203294 Sequence
	3	37	1.6	98	12	SYNPROCBA	M63418 Synthetic 5
	4	36.8	1.6	84	6	AR038854	AR038854 Sequence
	5	36.8	1.6	84	6	AR050387	AR050387 Sequence
	6	36.4	1.5	80	6	AR011130	AR011130 Sequence
	7	36.2	1.5	74	6	A08894	A08894 H.sapiens (
	8	36	1.5	77	6	A08908	A08908 H.sapiens (
	9	36	1.5	96	6	A08909	A08909 H.sapiens (
	10	35.8	1.5	85	6	BD003061	BD003061 Polynucle
	11	35.6	1.5	73	6	A08915	A08915 H.sapiens (
	12	35.4	1.5	85	6	AX099459	AX099459 Sequence
	13	35.4	1.5	97	6	AX260590	AX260590 Sequence
c	14	35.4	1.5	98	6	AX320854	AX320854 Sequence
	15	35	1.5	100	9	AF117819	AF117819 Homo sapi
	16	34.8	1.5	62	6	AX205289	AX205289 Sequence
c	17	34.8	1.5	100	6	AX260734	AX260734 Sequence
	18	34.6	1.5	80	6	A08907	A08907 H.sapiens (
	19	34.6	1.5	80	6	A18777	A18777 5' terminus
	20	34.6	1.5	80	6	AR001312	AR001312 Sequence
	21	34.6	1.5	80	6	AR130410	AR130410 Sequence
	22	34.6	1.5	84	6	A08920	A08920 H.sapiens (
c	23	34.6	1.5	85	6	AX182181	AX182181 Sequence
	24	34.6	1.5	91	3	DDIACTIN8A	M25215 Dictyosteli
	25	34.4	1.5	71	9	S76508	S76508 D1S8 (B) {h
	26	34.4	1.5	79	6	AX099452	AX099452 Sequence
	27	34.4	1.5	89	6	AX203299	AX203299 Sequence
	28	34.2	1.4	86	6	A08895	A08895 H.sapiens (
c	29	34.2	1.4	90	6	AR148129	AR148129 Sequence
c	30	34.2	1.4	90	6	E12580	E12580 Probe. 6/19
c	31	34	1.4	86	3	DDIACTC	M29111 D.discoideu
c	32	33.8	1.4	93	6	AX320853	AX320853 Sequence
	33	33.8	1.4	98	6	AX209704	AX209704 Sequence
c	34	33.8	1.4	100	10	RNU12531	U12531 Rattus norv
c	35	33.6	1.4	76	6	AX260960	AX260960 Sequence
c	36	33.6	1.4	79	6	AX182145	AX182145 Sequence
c	37	33.6	1.4	81	6	AX182108	AX182108 Sequence
	38	33.2	1.4	62	6	AX205290	AX205290 Sequence
	39	33.2	1.4	62	6	AX205291	AX205291 Sequence
c	40	33.2	1.4	70	6	AX260894	AX260894 Sequence
	41	33	1.4	69	6	BD005374	BD005374 Secreted
	42	33	1.4	69	6	BD005384	BD005384 Secreted
	43	33	1.4	69	6	I89934	I89934 Sequence 7
	44	33	1.4	69	6	I89944	I89944 Sequence 22
c	45	33	1.4	74	6	AX322391	AX322391 Sequence

ALIGNMENTS

RESULT 1
A08898
LOCUS A08898 99 bp DNA linear PAT 02-SEP-1993
DEFINITION H.sapiens (haplotype 2B, allele MS32, isolate French, serial number 16) minisatellite sequence.
ACCESSION A08898
VERSION A08898.1 GI:411820
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 99)
AUTHORS Jeffreys,A.J.
TITLE Extended nucleotide sequences
JOURNAL Patent: EP 0370719-A 81 30-MAY-1990;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES Location/Qualifiers
source 1. .99
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 67 a 0 c 0 g 32 t
ORIGIN

```
  Query Match             1.7%;  Score 39.4;  DB 6;  Length 99;
  Best Local Similarity  75.4%;  Pred. No. 3.4e+02;
  Matches   49;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   2301 ttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaa 2360
          ||||||| ||| ||||| |||||   || |  | |||  || | |||| | |||||||||
Db      2 TTTTATATTTTATTTAAAAAAAAATTAAATTATAAAAAATATATAAAATATAAAAAAAAA 61

Qy   2361 aaaaa 2365
          |||||
Db     62 AAAAA 66


RESULT   2
AX203294
LOCUS       AX203294                  86 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 39 from Patent WO0153836.
ACCESSION   AX203294
VERSION     AX203294.1  GI:15392662
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Schlegel,R., Endege,W. and Monahan,J.E.
  TITLE     Identification, assessment, prevention, and therapy of prostate
            cancer
  JOURNAL   Patent: WO 0153836-A 39 26-JUL-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source      -   1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       65 a      5 c      3 g     12 t      1 others
ORIGIN


  Query Match             1.7%;  Score 39.2;  DB 6;  Length 86;
  Best Local Similarity  66.7%;  Pred. No. 3.7e+02;
  Matches   56;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy   2282 gaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaactta 2341
          | |  | ||||||  | |     | ||  |  |||| ||||| | ||  |   |||   |
Db      2 GTACTTCTAGAATTAATTAAAACAGACAATTCTTAAAAAAAAAAAAAAAAAAAAAAAAAA 61

Qy   2342 aaaaaaaaaaaaaaaaaaaaaa 2365
          ||||||||||||||||||||||
Db     62 AAAAAAAAAAAAAAAAAAAAAA 85


RESULT   3
SYNPROCBA
LOCUS       SYNPROCBA                 98 bp    DNA     linear   SYN 27-APR-1993
DEFINITION  Synthetic 5' promoter region of C.biennis spheroidin gene.
ACCESSION   M63418
VERSION     M63418.1  GI:209140
KEYWORDS    .
SOURCE      Synthetic DNA.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Pearson,A., Richardson,C. and Yuen,L.
  TITLE     The 5' noncoding region sequence of the Choristoneura biennis
            Entomopoxvirus spheroidin gene funcions as an efficient late
            promoter in the mammalian vaccinia expression system
  JOURNAL   Virology 180, 561-566 (1991)
  MEDLINE   91111975
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       53 a      9 c      4 g     32 t
ORIGIN


  Query Match             1.6%;  Score 37;  DB 12;  Length 98;
  Best Local Similarity  64.7%;  Pred. No. 1.1e+03;
  Matches   55;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy   2280 ttgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaact 2339
          ||  || |    |  ||| | |||||      || ||||||||||||||  |||||
Db      5 TTATAACTTATAATACCAATATTTTACTACAACTCTAATAAAAATAGAATAATTTATTTA 64

Qy   2340 taaaaaaaaaaaaaaaaaaaaaaa 2364
          | | ||| ||  ||||| |||||||
Db     65 TTATAAATAAGCAAAAATAAAAAAA 89


RESULT   4
AR038854
LOCUS       AR038854                  84 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 3 from patent US 5807703.
ACCESSION   AR038854
VERSION     AR038854.1  GI:5958217
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Evans,C., Spaulding,V. and Bowman,M.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5807703-A 3 15-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="unknown"
BASE COUNT       76 a      0 c      0 g      7 t      1 others
ORIGIN


  Query Match             1.6%;  Score 36.8;  DB 6;  Length 84;
  Best Local Similarity  77.2%;  Pred. No. 1.2e+03;
  Matches   44;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2309 tttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
          ||| ||| | ||||| | ||  |   |||   |||||||||||||||||||||||||
Db      1 TTTTTTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 57


RESULT   5
AR050387
LOCUS       AR050387                  84 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 3 from patent US 5827688.
ACCESSION   AR050387
VERSION     AR050387.1  GI:5973112
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M. and Spaulding,V.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5827688-A 3 27-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="unknown"
BASE COUNT       76 a      0 c      0 g      7 t      1 others
ORIGIN


  Query Match             1.6%;  Score 36.8;  DB 6;  Length 84;
  Best Local Similarity  77.2%;  Pred. No. 1.2e+03;
```

```
  Matches   44;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   2309 tttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaa 2365
          ||| ||| | ||||| | ||  |   |||   ||||||||||||||||||||||||
Db      1 TTTTTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 57


RESULT   6
AR011130
LOCUS       AR011130                80 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 6 from patent US 5762924.
ACCESSION   AR011130
VERSION     AR011130.1  GI:3969120
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Dall,D.James., Fernon,C.Anne. and Sriskantha,A.
  TITLE     Recombinant entomopoxvirus
  JOURNAL   Patent: US 5762924-A 6 09-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="unknown"
BASE COUNT       46 a      8 c      2 g     24 t
ORIGIN


  Query Match             1.5%;  Score 36.4;  DB 6;  Length 80;
  Best Local Similarity   70.0%;  Pred. No. 1.5e+03;
  Matches   49;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy   2295 ccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaa 2354
           ||| | |||||      || ||||||||||||||  |||||    | | ||| ||  |||
Db      5 CCAATATTTTACTACAACTCTAATAAAAATAGAATAATTTATTTATTATAAATAAGCAAA 64

Qy   2355 aaaaaaaaaa 2364
          ||| ||||||
Db     65 AAATAAAAAA 74


RESULT   7
A08894
LOCUS       A08894                  74 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (haplotype 2A, allele MS32, isolate English, serial
            number 9) minisatellite sequence.
ACCESSION   A08894
VERSION     A08894.1  GI:411816
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 77 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       52 a      0 c      0 g     22 t
ORIGIN


  Query Match             1.5%;  Score 36.2;  DB 6;  Length 74;
  Best Local Similarity   72.3%;  Pred. No. 1.6e+03;
  Matches   47;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   2301 ttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaa 2360
           ||||||| ||| || || ||||| | ||   | ||||   || ||||  | |||||| ||
Db      2 TTTTATATTTTATTAAAAAAAAAAATAAAATTATAAAAAAAATAAAATTATAAAAAATAA 61

Qy   2361 aaaaa 2365
          |||||
Db     62 AAAAA 66


RESULT   8
A08908
LOCUS       A08908                  77 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (haplotype 3, allele MS32, isolate Mormon, serial number
            10) minisatellite sequence.
ACCESSION   A08908
VERSION     A08908.1  GI:411830
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 91 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       72 a      0 c      0 g      5 t
ORIGIN


  Query Match             1.5%;  Score 36;  DB 6;  Length 77;
  Best Local Similarity   80.8%;  Pred. No. 1.8e+03;
  Matches   42;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2314 ttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
           |||||||||| | ||  |   |||   ||||||||||||||||||||||||
Db      1 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 52


RESULT   9
A08909
LOCUS       A08909                  96 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (haplotype 3, allele MS32, isolate English, serial number
            15) minisatellite sequence.
ACCESSION   A08909
VERSION     A08909.1  GI:411831
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 92 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       91 a      0 c      0 g      5 t
ORIGIN


  Query Match             1.5%;  Score 36;  DB 6;  Length 96;
  Best Local Similarity   80.8%;  Pred. No. 1.8e+03;
  Matches   42;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   2314 ttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
           |||||||||| | ||  |   |||   ||||||||||||||||||||||||
Db      1 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 52
```

```
RESULT  10
BD003061
LOCUS       BD003061                  85 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Polynucleotide encoding secretory protein.
ACCESSION   BD003061
VERSION     BD003061.1  GI:18631022
KEYWORDS    JP 2001501455-A/3.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Evans,C., Bowman,M. and Spaulding,V.
  TITLE     Polynucleotide encoding secretory protein
  JOURNAL   Patent: JP 2001501455-A 3 06-FEB-2001;
            GENETICS INSTITUTE INC
COMMENT     OS   Unidentified
            PN   JP 2001501455-A/3
            PD   06-FEB-2001
            PF   06-JUN-1997 JP 1998500884
            PR   07-JUN-1996 JP    08/659224
            PI   KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
              DAVID MERBERG,
            PI   MAURICE TREACY,CHERYL EVANS,MICHAEL BOWMAN,VIKKI SPAULDING PC
             C12N15/09,A23K1/16,C07K14/47,C07K14/52,C12N5/10,C12N15/00, PC
            C12N5/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .85
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       80 a      0 c      0 g      5 t
ORIGIN
```

```
  Query Match             1.5%;  Score 35.8;  DB 6;  Length 85;
  Best Local Similarity   78.2%;  Pred. No. 2e+03;
  Matches   43;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   2311 tctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
          | ||||| ||||| | ||  |   |||   ||||||||||||||||||||||||
Db      1 TTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 55


RESULT  11
A08915
LOCUS       A08915                    73 bp    DNA     linear   PAT 02-SEP-1993
DEFINITION  H.sapiens (haplotype 3, allele MS32, isolate French, serial number
            8) minisatellite sequence.
ACCESSION   A08915
VERSION     A08915.1  GI:411837
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 98 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       69 a      0 c      0 g      4 t
ORIGIN
```

```
  Query Match             1.5%;  Score 35.6;  DB 6;  Length 73;
  Best Local Similarity   82.0%;  Pred. No. 2.2e+03;
  Matches   41;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   2316 aataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
          || ||||| | ||  |   ||| ||||||||||||||||||||||||||
Db     21 AAAAAAAAAAAAAAAAAAAAAAATTAAAAAAAAAAAAAAAAAAAAAAAA 70


RESULT  12
AX099459
LOCUS       AX099459                  85 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 99 from Patent WO0119988.
ACCESSION   AX099459
VERSION     AX099459.1  GI:13538559
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Jacobs,K., Mccoy,J.M., Lavallie,E.R., Collins-Racie,L.A., Evans,C.,
            Merberg,D., Treacy,M., Bowman,M.R., Spaulding,V. and Agostino,M.J.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: WO 0119988-A 99 22-MAR-2001;
            Genetics Institute, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .85
                     /note="n = A,T,C or G"
BASE COUNT       58 a      7 c      4 g     14 t      2 others
ORIGIN
```

```
  Query Match             1.5%;  Score 35.4;  DB 6;  Length 85;
  Best Local Similarity   64.6%;  Pred. No. 2.4e+03;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy   2287 tatagaatccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaa 2346
          || | || |||   |  | |  |    ||   ||||| | ||  |   |||   ||||||
Db      7 TAAATAAACCATGNTGATTTNTTAAACTTGCCAAAAAAAAAAAAAAAAAAAAAAAAAA 66

Qy   2347 aaaaaaaaaaaaaaaaaa 2365
          ||||||||||||||||||
Db     67 AAAAAAAAAAAAAAAAAA 85


RESULT  13
AX260590
LOCUS       AX260590                  97 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 241 from Patent WO0173027.
ACCESSION   AX260590
VERSION     AX260590.1  GI:16509557
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Meagher,M.J., Xu,J. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of colon cancer
  JOURNAL   Patent: WO 0173027-A 241 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       15 a     25 c     20 g     33 t      4 others
ORIGIN
```

```
  Query Match             1.5%;  Score 35.4;  DB 6;  Length 97;
  Best Local Similarity   94.7%;  Pred. No. 2.4e+03;
  Matches   36;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 agctcggatccctagtaacggccgccagtgtgctggaa 47
           ||||| |||||||||||||| ||||||||||||||||||
Db       1 AGCTCNGATCCCTAGTAACAGCCGCCAGTGTGCTGGAA 38


RESULT  14
AX320854/c
LOCUS       AX320854                  98 bp    DNA     linear   PAT 14-DEC-2001
DEFINITION  Sequence 24 from Patent WO0183736.
ACCESSION   AX320854
VERSION     AX320854.1  GI:17902405
KEYWORDS    .
SOURCE      Hepatitis C virus.
  ORGANISM  Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE   1  (sites)
  AUTHORS   Pellerin,C. and Kukolj,G.
  TITLE     Internal de novo initiation sites of the hcv ns5b polymerase and
            use thereof
  JOURNAL   Patent: WO 0183736-A 24 08-NOV-2001;
            BOEHRINGER INGELHEIM (CANADA) LTD. (CA)
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Hepatitis C virus"
                     /db_xref="taxon:11103"
BASE COUNT        2 a      5 c      1 g     90 t
ORIGIN


  Query Match             1.5%;  Score 35.4;  DB 6;  Length 98;
  Best Local Similarity   79.2%;  Pred. No. 2.4e+03;
  Matches   42;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    2313 tttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
           | ||| ||||| | ||| |   |||   ||||||||||||||||||||||||
Db      90 TATAAAAAAAAAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 38


RESULT  15
AF117819
LOCUS       AF117819                 100 bp    mRNA    linear   PRI 17-SEP-2001
DEFINITION  Homo sapiens bradykinin B1 receptor mRNA, partial cds and
            3'-untranslated region.
ACCESSION   AF117819
VERSION     AF117819.1  GI:4325046
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Zhou,X., Prado,G.N., Chai,M., Yang,X., Taylor,L. and Polgar,P.
  TITLE     Posttranscriptional destabilization of the bradykinin B1 receptor
            messenger RNA: cloning and functional characterization of the
            3'-untranslated region
  JOURNAL   Mol. Cell Biol. Res. Commun. 1 (1), 29-35 (1999)
  MEDLINE   20393235
   PUBMED   10329474
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Zhou,X. and Polgar,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-1999) Biochemistry, Boston University School of
            Medicine, 80 East Concord St., Boston, MA 02118, USA
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="IMR-90"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
                     /dev_stage="embryo"
     CDS             <1. .57
                     /note="BKB1R"
                     /codon_start=1
                     /product="bradykinin B1 receptor"
                     /protein_id="AAD17233.1"
                     /db_xref="GI:4325047"
                     /translation="APISSSHRKEIFQLFWRN"
     polyA_signal    53. .59
                     /note="alternative; overlaps with stop codon"
     3'UTR           58. .100
     polyA_site      72
BASE COUNT       53 a     18 c     10 g     19 t
ORIGIN


  Query Match             1.5%;  Score 35;  DB 9;  Length 100;
  Best Local Similarity   66.7%;  Pred. No. 3e+03;
  Matches   50;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy    2291 gaatccagtgttttatacttttctttaataaaaatagaagtatttaaacttaaaaaaaaaa 2350
           |||   | | ||  |   |||||     ||  | |  || ||| || |  ||||||||||
Db      24 GAAAGAAATCTTCCAACTTTTCTGGCGGAATTAAAACAGCATTGAACCAAAAAAAAAAAA 83

Qy    2351 aaaaaaaaaaaaaaa 2365
           |||||||||||||||
Db      84 AAAAAAAAAAAAAAA 98


Search completed: September 24, 2002, 21:07:56
Job time: 6918 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:14:08 ; Search time 370.16 Seconds
(without alignments)
10969.588 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 2046006

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		39.2	1.7	86	22	AAS23317	Human prostate can
2		38.6	1.6	95	13	AAQ28050	Entomopoxvirus sph
3		37.8	1.6	79	19	AAV26706	Human novel secret
4		36.8	1.6	84	19	AAV09269	Nucleotide sequenc
5		36.8	1.6	84	19	AAV04275	Secreted protein C
6	c	36.6	1.5	51	22	AAL32186	Human SNP oligonuc
7		36.6	1.5	99	18	AAT91300	Human M97-2 secret
8	c	36.4	1.5	92	18	AAT88439	Padlock probe alph
9	c	36.2	1.5	90	21	AAF18395	Lung cancer associ
10		36	1.5	53	22	ABA11098	Human nervous syst
11	c	36	1.5	90	12	AAQ11762	Self-complementary
12		36	1.5	94	16	AAQ89696	Vector back-end ol
13	c	35.8	1.5	80	19	AAV37197	Oligonucleotide se
14		35.8	1.5	85	19	AAV05720	Nucleotide sequenc
15	c	35.6	1.5	51	22	AAL29620	Human SNP oligonuc
16		35.4	1.5	85	19	AAV32418	Homo sapiens clone
17		35.4	1.5	85	22	AAF98434	Human cDNA clone B
18		35.4	1.5	97	23	AAS57565	cDNA #241 encoding
19	c	35.4	1.5	98	12	AAQ11764	Self-complementary
20		35	1.5	93	21	AAC98487	Human colon cancer
21	c	35	1.5	94	16	AAQ89695	Vector back-end ol
22		34.8	1.5	61	21	AAC24838	Human secreted pro
23		34.8	1.5	62	22	AAD13991	Deletion construct
24	c	34.8	1.5	100	23	AAS57709	cDNA #385 encoding
25		34.6	1.5	63	21	AAC14751	Human secreted pro
26		34.6	1.5	81	21	AAC13249	Human secreted pro
27	c	34.6	1.5	85	22	AAS07817	Cervical cancer pr
28		34.4	1.5	79	19	AAV32414	Homo sapiens clone
29		34.4	1.5	79	22	AAF98430	Human cDNA clone B
30	c	34.4	1.5	86	12	AAQ11760	Self-complementary
31		34.4	1.5	89	22	AAS23322	Human prostate can
32	c	34.4	1.5	98	20	AAX00174	Porcine reproducti
33	c	34.2	1.4	90	18	AAT47082	Synthetic DNA prob
34	c	34.2	1.4	90	19	AAV37789	Analytical solid p
35	c	34	1.4	51	22	AAL29656	Human SNP oligonuc
36	c	34	1.4	51	22	AAL32448	Human SNP oligonuc
37	c	34	1.4	77	22	AAS29173	Genomic sequence #
38		34	1.4	77	22	AAK78766	Human immune/haema
39		34	1.4	77	22	AAK79578	Human immune/haema
40		34	1.4	77	22	AAK84669	Human immune/haema
41		33.8	1.4	69	18	AAT88081	3' portion of cDNA
42		33.8	1.4	69	18	AAV02148	Human secreted pro
43		33.8	1.4	98	22	AAS25363	Human ovarian PCR-
44		33.6	1.4	60	21	AAC15704	Human secreted pro
45	c	33.6	1.4	76	23	AAS57935	cDNA #611 encoding

ALIGNMENTS

RESULT 1
AAS23317
ID AAS23317 standard; cDNA; 86 BP.
XX
AC AAS23317;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human prostate cancer marker #39.
XX
KW Prostate cancer; human; cancer marker; metastasis; diagnostic;
KW therapeutic; immunogenic; ss.
XX
OS Homo sapiens.
XX
PN WO200153836-A2.
XX
PD 26-JUL-2001.
XX
PF 24-JAN-2001; 2001WO-US02318.
XX
PR 24-JAN-2000; 2000US-0178525.
PR 17-FEB-2000; 2000US-0183245.
PR 16-MAR-2000; 2000US-0190139.
PR 31-MAY-2000; 2000US-0208126.
PR 18-JUL-2000; 2000US-0219705.
PR 13-DEC-2000; 2000US-0255160.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege W, Monahan JE;
XX

DR WPI; 2001-432912/46.
XX
PT Detecting and characterizing human prostate cancer comprises comparing
PT level of expression of marker in patient sample to normal level of
PT expression of marker in control non-prostate cancer sample -
XX
PS Claim 1; Page 874; 975pp; English.
XX
CC The invention relates to a method of assessing whether a patient is
CC afflicted with prostate cancer (PCA) comprising comparing the level of
CC expression of a marker in a patient sample to the normal level of marker
CC expression in a control non-PCA sample, where a significant difference
CC between marker expression in the patient sample and the normal level is
CC an indication that the patient is afflicted with PCA. The method can also
CC be used for: (1) monitoring the progression of PCA in a patient; (2)
CC assessing the efficacy of a test compound for inhibiting PCA in a
CC patient; (3) assessing the efficacy of a therapy for inhibiting PCA in a
CC patient; (4) selecting a composition for inhibiting PCA in a patient;
CC (5) inhibiting PCA in a patient comprising administering the selected
CC composition; (6) making an isolated hybridoma which produces an antibody
CC useful for carrying out (1) comprising immunising a mammal using the
CC marker, isolating splenocytes from the mammal, fusing the splenocytes
CC with an immortalised cell line and screening individual hybridomas for
CC production of an antibody which specifically binds the marker; (7) an
CC antibody produced by the hybridoma of (6); (8) assessing the prostate
CC cell carcinogenic potential of a test compound; (9) inhibiting PCA in a
CC patient at risk for developing PCA; and (10) determining whether PCA has
CC metastasised in a patient or assessing the aggressiveness or indolence
CC of PCA. AAS23279-AAS23379 represent the coding sequences of prostate
CC cancer markers used in the method of the invention.
XX
SQ Sequence 86 BP; 65 A; 5 C; 3 G; 12 T; 1 other;

```
  Query Match             1.7%;  Score 39.2;  DB 22;  Length 86;
  Best Local Similarity   66.7%;  Pred. No. 23;
  Matches   56;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy   2282 gaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaactta 2341
           | |  | ||||||  | |     | ||  |  |||| ||||| | ||  |   |||   |
Db      2 gtacttctagaattaattaaaacagacaattcttaaaaaaaaaaaaaaaaaaaaaaaaaa 61

Qy   2342 aaaaaaaaaaaaaaaaaaaaaaa 2365
           |||||||||||||||||||||||
Db     62 aaaaaaaaaaaaaaaaaaaaaaa 85
```

RESULT 2
AAQ28050
ID AAQ28050 standard; DNA; 95 BP.
XX
AC AAQ28050;
XX
DT 19-FEB-1993 (first entry)
XX
DE Entomopoxvirus spherodin promoter.
XX
KW Recombinant; expression vector; ss.
XX
OS Choristoneura biennis.
XX

FH	Key	Location/Qualifiers
FT	misc_feature	35..74
FT		/*tag= a
FT		/note= "essential promoter"

XX
PN CA2033957-A.
XX
PD 09-JUL-1992.
XX
PF 10-JAN-1991; 91CA-2033957.
XX
PR 08-JAN-1991; 91US-0638742.
XX
PA (YUEN/) YUEN K L.
XX
PI Arif B, Yuen KCL;
XX
DR WPI; 1992-316531/39.
XX
PT Recombinant vertebrate pox-virus expression vector - under
PT control of entomopoxvirus spheroidin gene promoter, allows for
PT efficient protein and vaccine prodn.
XX
PS Claim 17; Fig 3; 36pp; English.
XX
CC The sequence is that of the spheroidin promoter of Choristoneura
CC biennis. The promoter is used in the construction of a recombinant
CC vertebrate poxvirus which is capable of expressing a foreign protein
CC gene in a vertebrate tissue culture cell or in an animal susceptible
CC to the vaccinia virus. It may also be used in the prodn. of a
CC protein foreign to the poxvirus and as a vaccine consisting of at
CC least one immunogenic protein of a pathogen of a vertebrate animal
CC e.g. hepatitis B virus surface antigen, influenza virus haemagglutin,
CC herpes virus glycoprotein D, rabies virus glycoprotein and HIV
CC antigen.
XX
SQ Sequence 95 BP; 53 A; 9 C; 3 G; 30 T; 0 other;

```
  Query Match             1.6%;  Score 38.6;  DB 13;  Length 95;
  Best Local Similarity   65.9%;  Pred. No. 31;
  Matches   56;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy   2280 ttgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaact 2339
           ||  || |    |  ||| | |||||      || ||||||||||||||  |||||
Db      5 ttataacttataataccaatattttactacaactctaataaaaatagaataatttattta 64

Qy   2340 taaaaaaaaaaaaaaaaaaaaaaa 2364
          | | ||| ||  ||||||||||||
Db     65 ttataaataagcaaaaaaaaaaaa 89
```

RESULT 3
AAV26706
ID AAV26706 standard; cDNA; 79 BP.
XX
AC AAV26706;
XX
DT 11-SEP-1998 (first entry)
XX
DE Human novel secreted protein clone BF290_li 3'-end DNA.
XX
KW Secreted protein; human; marker; cytokine; immune stimulant; suppressor;
KW autoimmune disease; regulator; activin; inhibin; inhibitor; chemotactic;
KW chemokinetic; haemostatic; thrombocyte; tumour; anti-inflammatory; ds.
XX
OS Homo sapiens.
XX
PN WO9814470-A2.
XX
PD 09-APR-1998.
XX
PF 03-OCT-1997; 97WO-US18032.
XX
PR 04-OCT-1996; 96US-0725885.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
PI Spaulding V, Treacy M;
XX
DR WPI; 1998-240017/21.
XX
PT Compositions containing isolated proteins encoded by specific
PT nucleic acids - useful e.g. as anti-inflammatory, immune stimulant

```
PT   or suppressor, etc.
XX
PS   Disclosure; Page 94; 117pp; English.
XX
CC   This sequence encodes the 3'-end of a novel secreted protein from clone
CC   BF290_1i isolated from a human fetal brain cDNA library. Such secreted
CC   proteins can be used for, e.g. research purposes (as markers for
CC   tissues, molecular weight markers for gels, primers and probes), for
CC   nutrition (as C, N or carbohydrate source), as a cytokine for cell
CC   proliferation and differentiation activity, as immune stimulants or
CC   suppressors, e.g. for viral, bacterial or fungal infections, for
CC   autoimmune diseases such as multiple sclerosis or systemic lupus
CC   erythematosus, to regulate haematopoiesis, for tissue growth, as an
CC   activin or inhibin, or having chemotactic or chemo-kinetic, haemostatic
CC   and thrombocytic, receptor/ligand, anti-inflammatory or tumour inhibitory
CC   activities.
XX
SQ   Sequence 79 BP; 73 A; 0 C; 0 G; 6 T; 0 other;


  Query Match              1.6%;  Score 37.8;  DB 19;  Length 79;
  Best Local Similarity   78.9%;  Pred. No. 45;
  Matches   45;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    2309 tttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
           ||| ||||| ||||| | ||  |   |||   ||||||||||||||||||||||||
Db       1 tttatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 57


RESULT   4
AAV09269
ID   AAV09269 standard; cDNA; 84 BP.
XX
AC   AAV09269;
XX
DT   07-JUL-1998 (first entry)
XX
DE   Nucleotide sequence of the 3' portion of the C195_1 protein.
XX
KW   C195_1 protein; human PBMC cDNA library; cytokine activity; ss;
KW   peripheral blood mononuclear cell; nutritional activity; EST;
KW   cell proliferation/differentiation activity; expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO9748801-A2.
XX
PD   24-DEC-1997.
XX
PF   16-JUN-1997;   97WO-US10501.
XX
PR   17-JUN-1996;   96US-0664596.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Bowman M,  Evans C,  Jacobs K,  LaVallie ER,  McCoy JM;
PI   Merberg D,  Racie LA,  Spaulding V,  Treacy M;
XX
DR   WPI; 1998-063142/06.
XX
PT   Poly:nucleotide(s) and proteins obtained from human PBMC, dendritic
PT   cell, adult brain, foetal brain and adult testes cDNA libraries -
PT   used in research, detection and therapy of, e.g. cytokine and cell
PT   proliferation or differentiation
XX
PS   Claim 13; Page 48; 78pp; English.
XX
CC   This nucleotide sequence encodes the 3' portion of the C195_1 protein
CC   which was isolated from a human PBMC cDNA library.  The products of
CC   the polynucleotides of the invention can be used in research, detection
CC   and therapy, as they may have nutritional activity, cytokine and cell
CC   proliferation/differentiation activity.  A search against the Genbank
CC   database demonstrated that this sequence has at least some homology
CC   with two ESTs identified as "yg11g06.r1 Homo sapiens cDNA clone
CC   31937 5'" and "ym15f12.r1 Homo sapiens cDNA clone 48025 5'".
XX
SQ   Sequence 84 BP; 76 A; 0 C; 0 G; 7 T; 1 other;


  Query Match              1.6%;  Score 36.8;  DB 19;  Length 84;
  Best Local Similarity   77.2%;  Pred. No. 75;
  Matches   44;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    2309 tttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
           ||| ||| | ||||| | ||  |   |||   ||||||||||||||||||||||||
Db       1 ttttttttnaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 57


RESULT   5
AAV04275
ID   AAV04275 standard; cDNA; 84 BP.
XX
AC   AAV04275;
XX
DT   22-JUN-1998 (first entry)
XX
DE   Secreted protein C195_1 cDNA 3' end.
XX
KW   Secreted protein; C195_1; peripheral blood mononuclear cell; PBMC;
KW   protein factor; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9748800-A1.
XX
PD   24-DEC-1997.
XX
PF   16-JUN-1997;   97WO-US10500.
XX
PR   25-OCT-1996;   96US-0738367.
PR   17-JUN-1996;   96US-0664596.
PR   27-SEP-1996;   96US-0721926.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Lavallie ER,  McCoy JM,  Merberg D,  Racie LA;
PI   Spaulding V,  Treacy M;
XX
DR   WPI; 1998-063141/06.
XX
PT   Nucleic acid encoding secreted protein from human peripheral blood
PT   mononuclear cells - useful, e.g. as immunomodulators, antitumour
PT   agents, promoters of tissue growth, haemostatic and thrombolytic
PT   agents etc.
XX
PS   Disclosure; Page 36; 52pp; English.
XX
CC   This cDNA sequence comprises the 3' end, including the polyA tail,
CC   of cDNA clone C195_1 (see AAV04273), which codes for a novel secreted
CC   protein (see AAW41600) of human peripheral blood mononuclear cells
CC   (PBMC).  Clone C195_1 was isolated from a PBMC cDNA library using
CC   methods which are selective for cDNAs encoding secreted proteins.
CC   It is deposited as ATCC 98079.  C195_1 can be used to express
CC   recombinant secreted proteins; as a tissue/molecular weight marker;
CC   for chromosome identification; to identify possible genetic
CC   disorders; to isolate new related DNA; as a source of PCR primers;
CC   to generate anti-protein or anti-DNA antibodies; in interaction
CC   trap assays to identify sequences that encode interacting proteins
CC   etc.  The C195_1 secreted protein may have a number of biological
CC   activities.
XX
SQ   Sequence 84 BP; 76 A; 0 C; 0 G; 7 T; 1 other;


  Query Match              1.6%;  Score 36.8;  DB 19;  Length 84;
  Best Local Similarity   77.2%;  Pred. No. 75;
```

```
  Matches   44;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    2309 tttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
           ||| ||| | ||||| | ||  |   |||   |||||||||||||||||||||||||
Db       1 ttttttttnaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 57


RESULT   6
AAL32186/c
ID   AAL32186 standard; DNA; 51 BP.
XX
AC   AAL32186;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #5394.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2939; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 12 A; 0 C; 2 G; 37 T; 0 other;


  Query Match             1.5%;  Score 36.6;  DB 22;  Length 51;
  Best Local Similarity   82.4%;  Pred. No. 72;
  Matches   42;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    2315 taataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
           | |||||| |  || | | || | |||||||||||||||||||||||||
Db      51 TCATAAAATTTAAATTTTATACAATTAAAAAAAAAAAAAAAAAAAAAAAAA 1


RESULT   7
AAT91300
ID   AAT91300 standard; cDNA; 99 BP.
XX
AC   AAT91300;
XX
DT   19-MAY-1998  (first entry)
XX
DE   Human M97-2 secreted protein cDNA 3' end.
XX
KW   Secreted protein; anti-inflammatory; immune stimulant; suppressant;
KW   human; cytokine; autoimmune disease; regulator; activin; inhibin;
KW   M97-2; ds.
XX
OS   Homo sapiens.
XX
PN   WO9740151-A2.
XX
PD   30-OCT-1997.
XX
PF   14-APR-1997;   97WO-US06125.
XX
PR   10-JAN-1997;   97US-0781226.
PR   19-APR-1996;   96US-0635311.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Lavallie ER,  McCoy JM,  Merberg D,  Racie LA;
PI   Spaulding V;
XX
DR   WPI; 1997-535835/49.
XX
PT   Nucleic acids encoding novel secreted proteins - useful e.g. as an
PT   anti-inflammatory, immune stimulant or suppressor, etc.
XX
PS   Claim 16; Page 49; 81pp; English.
XX
CC   This sequence is the 3' end of a novel secreted protein, M97-2, which is
CC   isolated from a clone, M97-2 (ATCC 98028), derived from a human neural
CC   tissue (glioblastoma cell line T98G) cDNA library. The nucleic acid and
CC   encoded M97-2 secreted protein can be used for research purposes (as
CC   markers for tissues, molecular weight markers for gels, primers, probes,
CC   etc.), for nutrition (as C, N or carbohydrate source), as a cytokine for
CC   cell proliferation and differentiation activity, as immune stimulants or
CC   suppressors e.g. for viral, bacterial or fungal infections, for
CC   autoimmune diseases such as multiple sclerosis or systemic lupus
CC   erythematosus, to regulate haematopoiesis, for tissue growth, as an
CC   activin or inhibin, or having chemotactic, chemokinetic, haemostatic and
CC   thrombolytic, receptor/ligand, anti-inflammatory or tumour inhibitory
CC   activities.
XX
SQ   Sequence 99 BP; 86 A; 0 C; 2 G; 11 T; 0 other;


  Query Match             1.5%;  Score 36.6;  DB 18;  Length 99;
  Best Local Similarity   71.6%;  Pred. No. 87;
  Matches   48;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy    2299 tgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaa 2358
           | | || |  ||| |  || ||||| | ||  |   |||   |||||||||||||||||
Db       1 tatgttttgatttttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60

Qy    2359 aaaaaaa 2365
           |||||||
Db      61 aaaaaaa 67


RESULT   8
```

```
AAT88439/c
ID   AAT88439 standard; DNA; 92 BP.
XX
AC   AAT88439;
XX
DT   12-MAY-1998  (first entry)
XX
DE   Padlock probe alphaFV from WO9741254.
XX
KW   Cleavable padlock probe; target nucleic acid; detection;
KW   circularised structure; hybridisation; Factor V; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   67
FT                   /*tag=  a
FT                   /note= "attached to an amino group"
XX
PN   WO9741254-A1.
XX
PD   06-NOV-1997.
XX
PF   30-APR-1997;   97WO-SE00737.
XX
PR   30-APR-1996;   96SE-0001676.
XX
PA   (LAND/) LANDEGREN U.
XX
PI   Landegren U;
XX
DR   WPI; 1997-549751/50.
XX
PT   Detecting target nucleic acid sequence - using probe capable of
PT   forming circularised structure, which can be cleaved to form
PT   detectable function
XX
PS   Example 2; Fig 7; 20pp; English.
XX
CC   The present invention describes a method for detecting a target nucleic
CC   acid sequence (I) in a sample. The method comprises: (a) hybridising (I)
CC   to the probe ends of a probe having 2 free nucleic acid end parts, which
CC   are at least partially complementary to and capable of hybridising to
CC   at least 2 neighbouring regions of (I); (b) covalently connecting the
CC   ends of the hybridised probe with each other to form a circularised
CC   structure; (c) washing under denaturing conditions, characterised in
CC   that the probe is provided with a cleavable or dissociable detectable
CC   function; (d) cleaving or dissociating the detectable function; (e)
CC   separating probes with connected ends from probes with non-connected
CC   ends by washing under denaturing conditions; and (f) detecting the
CC   presence and, if desired, location of the remaining probe as indicative
CC   of the presence of (I). The present sequence represents a padlock probe,
CC   alpha FV, used in an example of the present invention. The method
CC   can be used to detect, quantify and distinguish between sequence
CC   variants with regard to 1 or several (I) in a sample. It can be used to
CC   distinguish between normal and mutated sequence variants associated with
CC   disease, for genetic linkage analysis of biallelic markers and to
CC   quantify gene expression in a tissue sample. The circularisable probe is
CC   designed so that it reports the presence of (I) by allowing a detectable
CC   moiety to remain bound only if the probe has been cyclised in a target
CC   dependent linking reaction. It also provides for high specificity and
CC   background reduction.
XX
SQ   Sequence 92 BP; 9 A; 9 C; 8 G; 66 T; 0 other;


  Query Match             1.5%;  Score 36.4;  DB 18;  Length 92;
  Best Local Similarity   70.0%;  Pred. No. 95;
  Matches   49;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy   2296 cagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaa 2355
           |||||  || || |    | || ||||| | ||  |   |||   || ||||||||||||
Db     92 CAGTGAATTGTAATACGATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAA 33

Qy   2356 aaaaaaaaaa 2365
           ||||||||||
Db     32 AAAAAAAAAA 23


RESULT   9
AAF18395/c
ID   AAF18395 standard; DNA; 90 BP.
XX
AC   AAF18395;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Lung cancer associated polynucleotide sequence SEQ ID 414.
XX
KW   Human; lung cancer associated protein; neuroprotective; cytostatic;
KW   cardioactive; immunomodulatory; muscular active; vulnerary;
KW   gastrointestinal; nephrotropic; antiinfective; gynecological;
KW   antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW   proliferative disorder; wound healing; infectious disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055180-A2.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05918.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Ruben SM;
XX
DR   WPI; 2000-587514/55.
DR   P-PSDB; AAB58519.
XX
PT   Lung cancer associated gene sequences, referred to as lung cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as lung cancer -
XX
PS   Claim 1; Page 883; 1425pp; English.
XX
CC   Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC   associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC   associated proteins and polynucleotide sequences, their agonists, and
CC   antagonists may have neuroprotective; cytostatic; cardioactive;
CC   immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC   general; nephrotropic; antiinfective; gynecological; or antibacterial
CC   activity. The invention also includes antibodies specific for the
CC   protein or polynucleotide sequences. The lung cancer associated
CC   polynucleotide sequences may be used for detection of lung cancer,
CC   chromosome identification, as chromosome markers, and for numerous other
CC   diagnostic or research purposes. The proteins may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders. The proteins may also be used in the treatment of wounds and
CC   infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC   peptide AAB58549 are used in the course of the invention for the
CC   identification and characterisation of the polynucleotide and protein
CC   sequences.
XX
SQ   Sequence 90 BP; 25 A; 4 C; 10 G; 46 T; 5 other;


  Query Match             1.5%;  Score 36.2;  DB 21;  Length 90;
  Best Local Similarity   70.1%;  Pred. No. 1e+02;
  Matches   47;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   2295 ccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaa 2354
```

```
          ||  | |||| |  | | |||  |||||| | ||  |   |||   ||||||||||||||
Db     81 CCNTTTTTTTTTTTTNTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 22

Qy   2355 aaaaaaa 2361
          |||||||
Db     21 AAAAAAA 15


RESULT  10
ABA11098
ID   ABA11098 standard; cDNA; 53 BP.
XX
AC   ABA11098;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 105.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   20-OCT-2000; 2000US-0242221.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
```

```
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250391.
PR   01-DEC-2000; 2000US-0251160.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-541565/60.
DR   P-PSDB; ABB14772.
XX
PT   Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating nervous system
PT   cancers and metastases -
XX
PS   Claim 1; SEQ ID NO 105; 1701pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC   (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 53 BP; 10 A; 17 C; 16 G; 10 T; 0 other;


  Query Match             1.5%;  Score 36;  DB 22;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 99;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 ctcggatccctagtaacggccgccagtgtgctggaa 47
           |||||||||||||||||||||||||||||||||||
Db      10 ctcggatccctagtaacggccgccagtgtgctggaa 45


RESULT  11
AAQ11762/c
ID   AAQ11762 standard; DNA; 90 BP.
XX
AC   AAQ11762;
XX
DT   22-JUL-1991  (first entry)
XX
DE   Self-complementary, T3 promoter hairpin-forming sequence #2.
XX
KW   T3 bacteriophage; DNA-dependent RNA polymerase;
KW   target sequence amplification; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       1..90
FT                   /*tag=  a
FT                   /note= "number of T's in loop is 50"
XX
PN   EP427073-A.
XX
PD   15-MAY-1991.
XX
PF   27-OCT-1990;   90EP-0120650.
XX
PR   23-AUG-1990;   90US-0569991.
PR   09-NOV-1989;   89US-0434372.
XX
PA   (MOLE-) MOLECULAR DIAGNOSTI.
XX
PI   Dattagupta N;
XX
DR   WPI; 1991-141638/20.
XX
PT   Nucleic acid probe for amplification and detection of target
PT   sequence - capable of forming ligatable hairpin structured
PT   promoter and transcribing target sequence, is sensitive and
PT   useful in medical diagnosis
XX
PS   Claim 2; Page 11; 15pp; English.
XX
CC   The sequence is an example of a preferred T3 RNA polymerase
CC   promoter for use in the invention. There can be from 2 to 50 T's in
CC   the loop region. A probe sequence is ligated to the 3' end of the
CC   promoter region. Upon hybridisation of the probe to a target sequence
CC   and ligation of the hybridised target sequence to the 5' end of the
CC   hairpin-forming sequence, the target sequence can be transcribed by
CC   T3 RNA polymerase. This allows the target sequence to be amplified
CC   (and detected) using a single oligonucleotide component (c.f. PCR).
CC   See also AAQ11759-Q11761 and AAQ11763-Q11764.
XX
SQ   Sequence 90 BP; 12 A; 8 C; 8 G; 62 T; 0 other;


  Query Match             1.5%;  Score 36;  DB 12;  Length 90;
  Best Local Similarity   80.8%;  Pred. No. 1.1e+02;
  Matches   42;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    2314 ttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
           |||||||||| | ||  |   |||   ||||||||||||||||||||||||
Db      75 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 24


RESULT  12
AAQ89696
ID   AAQ89696 standard; DNA; 94 BP.
XX
AC   AAQ89696;
XX
DT   07-NOV-1995  (first entry)
XX
DE   Vector back-end oligonucleotide (complementary strand).
XX
KW   Universal cloning vector; gene expression; cDNA expression;
KW   back-end oligonucleotide; ss.
XX
```

OS Synthetic.
XX
PN WO9510620-A.
XX
PD 20-APR-1995.
XX
PF 14-OCT-1994; 94WO-US11719.
XX
PR 15-OCT-1993; 93US-0136148.
XX
PA (MLCW) MALLINCKRODT VETERINARY INC.
XX
PI McMullen JR, Synenki RM, Zook CA;
XX
DR WPI; 1995-161805/21.
XX
PT Universal cloning vector for expression of heterologous genes or
PT cDNA - allows high levels of expression without the need for
PT sequence modification
XX
PS Claim 9; Fig.3; 40pp; English.
XX
CC A universal vector comprises a PL promoter, a ribosome binding site,
CC the first 30 bp of the delta-7 pig somatostatin gene, a 21 bp front-
CC end oligonucleotide (pref. the dimer comprising the complementary
CC strands given in AAQ89693-94) containing multiple blunt end restriction
CC enzyme sites, a back-end oligonucleotide (pref. the dimer
CC comprising the strands given in AAQ89695-96) containing at least 1
CC restriction enzyme site, a transcription terminator (pref. the dimer
CC comprising the strands given in AAQ89697-98) and a drug resistance
CC marker.
XX
SQ Sequence 94 BP; 20 A; 28 C; 26 G; 20 T; 0 other;

Query Match 1.5%; Score 36; DB 16; Length 94;
Best Local Similarity 97.9%; Pred. No. 1.2e+02;
Matches 47; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 ttggtaccgagctcggatcc-ctagtaacggccgccagtgtgctggaa 47
||||||||||||||||||| |||||||||||||||||||||||||||||
Db 4 ttggtaccgagctcggatccactagtaacggccgccagtgtgctggaa 51

RESULT 13
AAV37197/c
ID AAV37197 standard; DNA; 80 BP.
XX
AC AAV37197;
XX
DT 14-SEP-1998 (first entry)
XX
DE Oligonucleotide sequence of the specification.
XX
KW Genotype; phenotype; molecular evolutionary engineering;
KW functional biopolymer; virus; ss.
XX
OS Synthetic.
XX
PN WO9816636-A1.
XX
PD 23-APR-1998.
XX
PF 17-OCT-1997; 97WO-JP03766.
XX
PR 17-OCT-1996; 96JP-0274855.
XX
PA (MITU) MITSUBISHI CHEM CORP.
XX
PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
XX
DR WPI; 1998-261039/23.
XX
PT Virus containing nucleic acid and protein sections - for use in
PT modification and creation of functional bio:polymers such as
PT enzymes, antibodies and ribozyme(s)
XX
PS Example 1; Page 44; 68pp; Japanese.
XX
CC The present sequence is used in the course of the invention. The
CC specification describes a molecule for bringing together genotype with
CC phenotype (in vitro virus). The molecule contains a nucleic acid
CC fragment having a base sequence corresponding to a genotype, covalently
CC bonded to a protein fragment containing a protein participating in
CC phenotype expression, the 3'-end of the nucleic acid part being bonded
CC to the C-terminus of the protein part via a puromycin moiety. The nucleic
CC acid fragment preferably consists of RNA corresponding to the gene (free
CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
CC adapter (such as a DNA-tRNA hybrid containing an anticodon corresponding
CC to the terminator codon of the gene) and a puromycin-containing cap
CC capable of binding to an amino acid residue. Translation of the virus is
CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
CC ribosome), the protein synthesised by the translation attaching to the
CC puromycin cap, resulting in the complete in vitro virus structure. The
CC method is used in molecular evolutionary engineering to optimise function
CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
CC to generate new functionality. The virus can be used for the functional
CC optimisation of nucleic acid or protein sequences.
XX
SQ Sequence 80 BP; 7 A; 10 C; 1 G; 62 T; 0 other;

Query Match 1.5%; Score 35.8; DB 19; Length 80;
Best Local Similarity 78.2%; Pred. No. 1.2e+02;
Matches 43; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2311 tctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaa 2365
|| |||| ||||| | || | ||| |||||||||||||||||||||||
Db 68 TCATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14

RESULT 14
AAV05720
ID AAV05720 standard; cDNA; 85 BP.
XX
AC AAV05720;
XX
DT 05-JUN-1998 (first entry)
XX
DE Nucleotide sequence of the 3' portion from clone AZ302_1.
XX
KW Colon; secreted protein; EST sequence; homology; antibody;
KW immunoassay reagent; nutritional supplement; therapeutic activity; ds.
XX
OS Homo sapiens.
XX
PN WO9746683-A2.
XX
PD 11-DEC-1997.
XX
PF 06-JUN-1997; 97WO-US09878.
XX
PR 07-JUN-1996; 96US-0659224.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM;
PI Merberg D, Racie LA, Spaulding V, Treacy M;
XX
DR WPI; 1998-042191/04.
XX
PT Nucleic acids encoding secreted proteins from clones within ATCC
PT 98076 - useful as immuno-modulators, anti-proliferative agents,
PT regulators of cell differentiation and tissue growth, etc
XX
PS Claim 13; Page 57; 99pp; English.

XX
CC The present sequence represents the nucleotide sequence of the 3'
CC portion of clone AZ302_1. The clone was isolated from a human colon
CC (Caca-2 adenocarcinoma) cDNA library using probe AAV05749. AZ302_1 is
CC a full length clone encoding a secreted protein. AAV05719 provides the
CC 5' portion sequence and AAV05718 the internal sequence. The AZ302_1
CC sequence shows some homology with an EST sequence. The nucleic acid
CC can be used for expression of recombinant proteins, as tissue, molecular
CC weight or chromosome markers, indicators of genetic disorders and
CC sources of probes and primers. They can also be used to generate
CC anti-protein or anti-DNA antibodies and as components of interaction trap
CC assays etc. The protein is useful for raising antibodies, as immunoassay
CC reagents and as nutritional supplements. The protein may possibly have
CC any of a great variety of therapeutic activities.
XX
SQ Sequence 85 BP; 80 A; 0 C; 0 G; 5 T; 0 other;

Query Match 1.5%; Score 35.8; DB 19; Length 85;
Best Local Similarity 78.2%; Pred. No. 1.3e+02;
Matches 43; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2311 tctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
| ||||| ||||| | || | ||| ||||||||||||||||||||||||
Db 1 ttttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 55

RESULT 15
AAL29620/c
ID AAL29620 standard; DNA; 51 BP.
XX
AC AAL29620;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2828.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2195; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 10 A; 4 C; 2 G; 35 T; 0 other;

Query Match 1.5%; Score 35.6; DB 22; Length 51;
Best Local Similarity 82.0%; Pred. No. 1.2e+02;
Matches 41; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2316 aataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaaaaaaa 2365
|||||||| | | |||| ||||||||||||||||||||||||
Db 51 AATAAAAAGATTGGGATTTCCTTTTAAAAAAAAAAAAAAAAAAAAAAAAA 2

Search completed: September 24, 2002, 21:14:19
Job time: 7211 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:08:58 ; Search time 2466.96 Seconds
(without alignments)
12939.123 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttggtaccgagctcggatcc.....aaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 297742

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	48.2	2.0	79	9	AA781190	AA781190 aj24f04.s
c	2	44.4	1.9	86	9	AI701097	AI701097 wc78h09.x
	3	43	1.8	91	10	C25755	C25755 C25755 Dict
	4	42.6	1.8	96	10	R28767	R28767 F0-233D 22
c	5	42	1.8	83	9	AI499570	AI499570 to02g10.x
	6	42	1.8	89	9	AU053810	AU053810 AU053810
	7	42	1.8	93	9	AI945829	AI945829 bs17c12.y
	8	41.8	1.8	90	9	AU054081	AU054081 AU054081
	9	41.6	1.8	96	10	BI941603	BI941603 sc77d05.y
c	10	41.4	1.8	75	9	AI462919	AI462919 vb82g08.x
	11	41.4	1.8	94	9	AA241234	AA241234 mu95g06.r
	12	41.4	1.8	99	9	AU051957	AU051957 AU051957
c	13	41	1.7	86	9	AI636507	AI636507 ts91b09.x
c	14	41	1.7	93	10	BG122005	BG122005 602351665
c	15	41	1.7	98	10	BE319731	BE319731 NF020E07R
	16	40.6	1.7	96	9	AA120449	AA120449 mn47e11.r
c	17	40.4	1.7	92	9	AA554497	AA554497 ni36d04.s
c	18	40.4	1.7	99	9	AI638644	AI638644 tt25b02.x
	19	40.4	1.7	100	9	AI210082	AI210082 g9c01a1.r
c	20	40.2	1.7	73	9	AI050700	AI050700 ov42d11.s
c	21	40.2	1.7	87	9	AW085350	AW085350 wy62e08.x
	22	40.2	1.7	95	10	BF338778	BF338778 602036207
	23	40.2	1.7	100	9	AA607932	AA607932 vm39e04.r
c	24	39.8	1.7	73	10	R43518	R43518 yg19b08.s1
c	25	39.8	1.7	97	9	AI628325	AI628325 ts36f01.x
	26	39.6	1.7	100	9	AA624410	AA624410 vn03g08.r
c	27	39.4	1.7	64	9	AI955977	AI955977 wt32f10.x
c	28	39.4	1.7	68	10	N21402	N21402 yx54f07.s1
	29	39.4	1.7	80	10	BI142670	BI142670 kt44b10.y
c	30	39.4	1.7	83	9	AI061323	AI061323 an32h12.x
c	31	39.4	1.7	95	9	AI678446	AI678446 tu82e01.x
	32	39.4	1.7	97	9	AU038741	AU038741 AU038741
c	33	39.4	1.7	100	9	AI355779	AI355779 qt94d10.x
	34	39.2	1.7	75	9	AI938927	AI938927 sc63a02.y
	35	39.2	1.7	80	12	CNS01TV2	AL167015 Tetraodon
	36	39.2	1.7	97	10	C84172	C84172 C84172 Dict
	37	39	1.6	83	9	AU052639	AU052639 AU052639
c	38	39	1.6	87	9	AI419826	AI419826 tg52b06.x
c	39	39	1.6	88	10	BF635188	BF635188 NF081D04D
c	40	39	1.6	90	9	AI439664	AI439664 tc91h01.x
	41	39	1.6	93	9	AU038793	AU038793 AU038793
c	42	39	1.6	98	9	AW686658	AW686658 NF040D12N
c	43	39	1.6	99	9	AI611720	AI611720 tw67g09.x
c	44	38.8	1.6	72	9	AW148882	AW148882 xf05h02.x
c	45	38.8	1.6	78	2	HSM010524	A1045674 Homo sapi

ALIGNMENTS

RESULT 1
AA781190/c
LOCUS AA781190 79 bp mRNA linear EST 31-DEC-1998
DEFINITION aj24f04.s1 Soares_testis_NHT Homo sapiens cDNA clone 1391263 3' similar to TR:P96614 P96614 ATP-DEPENDENT RNA HELICASE DEAD HOMOLOG. . ;, mRNA sequence.
ACCESSION AA781190
VERSION AA781190.1 GI:2840521
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 79)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo , Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert Length: 708 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .79
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391263"
/clone_lib="Soares_testis_NHT"
/sex="male"

/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 25 a 17 c 13 g 24 t
ORIGIN

Query Match 2.0%; Score 48.2; DB 9; Length 79;
Best Local Similarity 76.6%; Pred. No. 1.1e+03;
Matches 59; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy   1853 gaggacttgatgtccatgacgttacacatgtctataatttgactttccacggaatattg 1912
          |||| |||||| |   ||| || |||||||| |||||| | ||  | ||| |||||||||
Db     78 GAGGTCTTGATCTTATTGATGTCACACATGTATATAATATGGATCTCCCAGGGAATATTG 19

Qy   1913 aagaatacgtacaccga 1929
          | | ||| |||||| ||
Db     18 ACGTATATGTACACAGA 2
```

RESULT 2
AI701097/c
LOCUS AI701097 86 bp mRNA linear EST 17-DEC-1999
DEFINITION wc78h09.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2324801 3' similar to contains TAR1.b2 MSR1 repetitive element ;, mRNA sequence.
ACCESSION AI701097
VERSION AI701097.1 GI:4988997
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 86)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1789 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 69.
FEATURES Location/Qualifiers
source 1. .86
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2324801"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"
BASE COUNT 17 a 3 c 2 g 64 t
ORIGIN

Query Match 1.9%; Score 44.4; DB 9; Length 86;
Best Local Similarity 73.1%; Pred. No. 3.8e+03;
Matches 57; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy   2288 atagaatccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaa 2347
          | | || |  || |||| |  ||| ||||| ||||| | ||  |   |||  |||||||
Db     85 AAAAAACCGGGTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 26

Qy   2348 aaaaaaaaaaaaaaaaa 2365
          |||||||||||||||||
Db     25 AAAAAAAAAAAAAAAAA 8
```

RESULT 3
C25755
LOCUS C25755 91 bp mRNA linear EST 29-JUL-1997
DEFINITION C25755 Dictyostelium discoideum FC (H.Urushihara) Dictyostelium discoideum cDNA clone FC-AZ08, mRNA sequence.
ACCESSION C25755
VERSION C25755.1 GI:2282541
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 91)
AUTHORS Suzuki,K., Shimizu,H. and Urushiihara,H.
TITLE Sexual cDNA in D. discoideum(970724)
JOURNAL Unpublished (1997)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402hu@sakura.cc.tsukuba.ac.jp.
FEATURES Location/Qualifiers
source 1. .91
/organism="Dictyostelium discoideum"
/strain="KAX3"
/db_xref="taxon:44689"
/clone="FC-AZ08"
/clone_lib="Dictyostelium discoideum FC (H.Urushihara)"
BASE COUNT 59 a 0 c 5 g 27 t
ORIGIN

Query Match 1.8%; Score 43; DB 10; Length 91;
Best Local Similarity 73.3%; Pred. No. 6e+03;
Matches 55; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy   2291 gaatccagtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaa 2350
          || |   || |||| |  ||| ||||| ||||| | ||  |   |||   ||||||||||
Db     10 GATTGTTGTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 69

Qy   2351 aaaaaaaaaaaaaaa 2365
          |||||||||||||||
Db     70 AAAAAAAAAAAAAAA 84
```

RESULT 4
R28767
LOCUS R28767 96 bp mRNA linear EST 03-DEC-1999
DEFINITION F0-233D 22 week old human fetal liver cDNA library Homo sapiens cDNA clone F0-233D 5', mRNA sequence.
ACCESSION R28767
VERSION R28767.1 GI:6514139
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 96)
AUTHORS Choi,S.S., Yun,J.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.S.
TITLE Construction of a gene expression profile of a human fetal liver by single-pass cDNA sequencing
JOURNAL Mamm. Genome 6 (9), 653-657 (1995)

```
  MEDLINE   96081342
COMMENT     Contact: Hee-Sup Shin
            Developmental Genetics
            Pohang Institute of Science & Technology
            San31, Hyojadong Pohang, 790-784 Republic of Korea
            Tel: 562-279-2291
            Fax: 562-279-2199
            Email: shinhs@vision.postech.ac.kr
            Seq primer: T3 primer.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="F0-233D"
                     /clone_lib="22 week old human fetal liver cDNA library"
                     /lab_host="XL1-blue MRF'"
                     /note="Vector: pBluescriptII SK(-); Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library made by oligo-dT primed and
                     directionally cloned between 5'ExoR I-XhoI3' sites."
BASE COUNT       27 a     19 c     22 g     28 t
ORIGIN


  Query Match             1.8%;  Score 42.6;  DB 10;  Length 96;
  Best Local Similarity   67.4%;  Pred. No. 6.8e+03;
  Matches   60;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy   1813 ggcaaagtgagaatactaattgcaactgatctagcctctagaggacttgatgtccatgac 1872
           || ||||     || || ||||| || ||| | ||||| ||||| || |||||  | ||
Db      2 GGAAAAGCTCCTATTCTGATTGCTACAGATGTGGCCTCCAGAGGGCTAGATGTGGAAGAT 61

Qy   1873 gttacacatgtctataattttgactttcc 1901
           || | |  ||||   |||| |||||  ||
Db     62 GTGAAATTTGTCATCAATTATGACTACCC 90


RESULT   5
AI499570/c
LOCUS       AI499570                  83 bp    mRNA    linear   EST 14-APR-1999
DEFINITION  to02g10.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2177922 3'
            similar to contains element MER26 repetitive element ;, mRNA
            sequence.
ACCESSION   AI499570
VERSION     AI499570.1  GI:4391552
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 887   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 67.
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2177922"
                     /clone_lib="NCI_CGAP_Ut2"
                     /tissue_type="moderately-differentiated endometrial
                     adenocarcinoma, 3 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.85 kb. Life Technologies catalog #:
                     11539-012"
BASE COUNT       17 a      1 c      3 g     62 t
ORIGIN


  Query Match             1.8%;  Score 42;  DB 9;  Length 83;
  Best Local Similarity   77.3%;  Pred. No. 8.8e+03;
  Matches   51;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2300 gttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaa 2359
           ||||| |  ||| ||||| ||||| | ||  |   |||   ||||||||||||||||||
Db     74 GTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 15

Qy   2360 aaaaaa 2365
           ||||||
Db     14 AAAAAA 9


RESULT   6
AU053810
LOCUS       AU053810                  89 bp    mRNA    linear   EST 28-APR-1999
DEFINITION  AU053810 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
            discoideum cDNA clone SLJ729, mRNA sequence.
ACCESSION   AU053810
VERSION     AU053810.1  GI:4702292
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     Developmental cDNA in Dictyostelium discoideum
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = Dictyostelium discoideum cDNA project in Japan.
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SLJ729"
                     /clone_lib="Dictyostelium discoideum SL (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       54 a      2 c      2 g     27 t      4 others
ORIGIN


  Query Match             1.8%;  Score 42;  DB 9;  Length 89;
  Best Local Similarity   66.3%;  Pred. No. 8.5e+03;
  Matches   57;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy   2280 ttgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaact 2339
           |||||||   |     |    ||||||| ||  || |||||   ||  | || |   |
Db      3 TTGAAAAANGAATTATCTAATTTTTATATTTATTTAAATAATTNTAATANTAATANTAAA 62

Qy   2340 taaaaaaaaaaaaaaaaaaaaaaaa 2365
           ||||||||||||||||||||||||
Db     63 AAAAAAAAAAAAAAAAAAAAAAAA 88


RESULT   7
```

```
AI945829
LOCUS       AI945829                 93 bp    mRNA    linear   EST 08-JAN-2001
DEFINITION  bs17c12.y1 Drosophila melanogaster adult testis library Drosophila
            melanogaster cDNA clone bs17c12 5', mRNA sequence.
ACCESSION   AI945829
VERSION     AI945829.1  GI:5736227
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Andrews,J., Bouffard,G.G., Cheadle,C., Lu,J., Becker,K.G. and
            Oliver,B.
  TITLE     Gene discovery using computational and microarray analysis of
            transcription in the drosophila melanogaster testis
  JOURNAL   Genome Res. 10 (12), 2030-2043 (2000)
  MEDLINE   20568492
COMMENT     Contact: Brian Oliver
            Laboratory of Cellular and Developmental Biology
            NIDDK, National Institutes of Health
            6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
            Fax: (301) 496 5239
            Email: oliver@helix.nih.gov,
            http://www.niddk.nih.gov/intram/people/boliver.htm
            Tissue isolation and library construction performed at the National
            Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
            http://www.niddk.nih.gov/intram/people/boliver.htm). DNA sequencing
            and analyses performed by National Institutes of Health Intramural
            Sequencing Center (NISC; see http://www.nisc.nih.gov).
            Plate: 17  row: c  column: 12
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Drosophila melanogaster"
                     /strain="y[*] w[67c1]/Y"
                     /db_xref="taxon:7227"
                     /clone="bs17c12"
                     /clone_lib="Drosophila melanogaster adult testis library"
                     /sex="male"
                     /dev_stage="1-5 day adult"
                     /lab_host="SOLR (Stratagene)"
                     /note="Organ: testis; Vector: pBlueScript SK (Stratagene);
                     Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5
                     day adult y[*] w[67c1]/Y males raised at 25oC.  RNA
                     isolated using Trizol (Life Technologies) and a single
                     round of Poly(A)+ selection using Oligotex (Qiagen).  cDNA
                     library constructed using Stratagene ZAP-cDNA syntesis
                     kit.  Oligo dT-primed, size fractionated ~1-6 kb, and
                     directionally cloned at EcoRI and XhoI in Uni-ZAP XR.
                     Following a single round of amplification pBlueScript SK
                     phagemids were mass excised.  A distribution channel for
                     clones is being sought, but not currently available.
                     Requests for clones cannot be honored."
BASE COUNT       63 a      6 c      8 g     16 t
ORIGIN


  Query Match             1.8%;  Score 42;  DB 9;  Length 93;
  Best Local Similarity   69.5%;  Pred. No. 8.4e+03;
  Matches   57;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy   2279 tttgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaac 2338
           | ||||||   | | || |   ||||| | |||||| | ||||| | ||  |   |||
Db     12 TCTGAAAAATAAAATTCAACGTTTTTAAAGTTTCTTAGAAAAAAAAAAAAAAAAAAAAAA 71

Qy   2339 ttaaaaaaaaaaaaaaaaaaa 2360
           |||||||||||||||||||
Db     72 AAAAAAAAAAAAAAAAAAAAA 93


RESULT   8
```

```
AU054081
LOCUS       AU054081                 90 bp    mRNA    linear   EST 28-APR-1999
DEFINITION  AU054081 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
            discoideum cDNA clone SLK564, mRNA sequence.
ACCESSION   AU054081
VERSION     AU054081.1  GI:4702562
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     Developmental cDNA in Dictyostelium discoideum
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = Dictyostelium discoideum cDNA project in Japan.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SLK564"
                     /clone_lib="Dictyostelium discoideum SL (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       63 a      3 c      2 g     22 t
ORIGIN


  Query Match             1.8%;  Score 41.8;  DB 9;  Length 90;
  Best Local Similarity   68.2%;  Pred. No. 9.1e+03;
  Matches   58;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy   2281 tgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaaactt 2340
           | || | |||   |  | | | ||  |  || | | ||||||| | ||  | | |||
Db      5 TTAATAAATATGTTGTATTTTATTTAAACTTATATTATAAAAAAAAAAACAATAAAAAAA 64

Qy   2341 aaaaaaaaaaaaaaaaaaaaaaaaa 2365
           |||||||||||||||||||||||||
Db     65 AAAAAAAAAAAAAAAAAAAAAAAAA 89


RESULT   9
BI941603
LOCUS       BI941603                 96 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  sc77d05.y1 Gm-c1018 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1018-898 5', mRNA sequence.
ACCESSION   BI941603
VERSION     BI941603.1  GI:16275960
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
```

```
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1018-898"
                     /clone_lib="Gm-c1018"
                     /tissue_type="leaves of greenhouse grown plants"
                     /dev_stage="2-3 weeks old"
                     /lab_host="DH10B (Gibco BRL)"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from 2-3 week old greenhouse grown plants.The cDNA library
                     was prepared using the Stratagene pBluescript II XR
                     library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into DH10B host cells (Gibco BRL). This
                     library was constructed by Dr. Randy Shoemaker and Dr.
                     John Erpelding."
BASE COUNT       66 a      6 c     10 g     14 t
ORIGIN


  Query Match             1.8%;  Score 41.6;  DB 10;  Length 96;
  Best Local Similarity   67.0%;  Pred. No. 9.5e+03;
  Matches   59;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy    2277 tgtttgaaaatatagaatccagtgttttatactttctttaataaaaatagaagtatttaa 2336
           | | | |||||  |||||      |    |||||| |  || ||||| | ||  |   ||
Db       9 TATCTAAAAATGAAGAATTATTCATGAAGTACTTTGTGGAAAAAAAAAAAAAAAAAAAAA 68

Qy    2337 acttaaaaaaaaaaaaaaaaaaaaaaa 2364
           |   |||||||||||||||||||||||
Db      69 AAAAAAAAAAAAAAAAAAAAAAAAAAA 96


RESULT  10
AI462919/c
LOCUS       AI462919                  75 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  vb82g08.x1 Soares mouse 3NME12 5 Mus musculus cDNA clone
            IMAGE:763550 3', mRNA sequence.
ACCESSION   AI462919
VERSION     AI462919.1  GI:4316949
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:464470
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            High quality sequence stop: 63.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:763550"
                     /clone_lib="Soares mouse 3NME12 5"
                     /sex="unknown"
                     /tissue_type="fetus"
                     /dev_stage="12.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCTTATTTTTTTTTTTTTTTT
                     3'],  on total  mouse RNA [provided by Minoru Ko, Wayne
                     State Univ.]; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       20 a      0 c      0 g     55 t
ORIGIN


  Query Match             1.8%;  Score 41.4;  DB 9;  Length 75;
  Best Local Similarity   76.1%;  Pred. No. 1.1e+04;
  Matches   51;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2299 tgttttatacttttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaa 2358
           | |||| |  ||| ||||| ||||| | ||  |   |||   ||||||||||||||||||
Db      72 TTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 13

Qy    2359 aaaaaaa 2365
           |||||||
Db      12 AAAAAAA 6


RESULT  11
AA241234
LOCUS       AA241234                  94 bp    mRNA    linear   EST 04-MAR-1997
DEFINITION  mu95g06.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
            IMAGE:653338 5', mRNA sequence.
ACCESSION   AA241234
VERSION     AA241234.1  GI:1864568
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
```

```
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          MGI:399186
          Putative full length read
          vector to vector length is 117
          Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:653338"
                     /clone_lib="Soares mouse lymph node NbMLN"
                     /sex="male"
                     /tissue_type="lymph node"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGATACTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library constructed and
                     normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       46 a     10 c     15 g     23 t
ORIGIN


  Query Match             1.8%;  Score 41.4;  DB 9;  Length 94;
  Best Local Similarity   76.1%;  Pred. No. 1e+04;
  Matches   51;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2298 gtgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaa 2357
           ||| ||    ||||  |||||||| |||  || | | |||   |||||||||||||||||
Db      28 GTGGTTGTCCCTTTTGTTAATAAACATATGAGCAGTCAAAAAAAAAAAAAAAAAAAAAAA 87

Qy    2358 aaaaaaa 2364
           |||||||
Db      88 AAAAAAA 94


RESULT  12
AU051957
LOCUS       AU051957                 99 bp    mRNA    linear   EST 26-MAR-1999
DEFINITION  AU051957 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA, mRNA sequence.
ACCESSION   AU051957
VERSION     AU051957.1  GI:4526898
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     Developmental cDNA in Dictyostelium discoideum
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT       73 a      1 c      2 g     23 t
ORIGIN


  Query Match             1.8%;  Score 41.4;  DB 9;  Length 99;
  Best Local Similarity   76.1%;  Pred. No. 1e+04;
  Matches   51;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    2299 tgttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaa 2358
           | |||| |  ||| ||||| ||||| | ||  |   |||   |||||||||||||||||
Db      16 TTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 75

Qy    2359 aaaaaaa 2365
           |||||||
Db      76 AAAAAAA 82


RESULT  13
AI636507/c
LOCUS       AI636507                 86 bp    mRNA    linear   EST 26-APR-1999
DEFINITION  ts91b09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2238617 3',
            mRNA sequence.
ACCESSION   AI636507
VERSION     AI636507.1  GI:4687837
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2238617"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
                     from the normalized library NCI_CGAP_GC4 was prepared, and
                     ss circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     1257096-1258631, 1469064-1470983, and 1475592-1476743).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       12 a      6 c      9 g     59 t
ORIGIN


  Query Match             1.7%;  Score 41;  DB 9;  Length 86;
  Best Local Similarity   76.9%;  Pred. No. 1.2e+04;
  Matches   50;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    2301 ttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaa 2360
```

```
          ||||   |||| ||||| ||||| | ||  |   |||   ||||||||||||||||||||
Db     69 TTTTGGGCTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 10

Qy   2361 aaaaa 2365
          |||||
Db      9 AAAAA 5


RESULT  14
BG122005/c
LOCUS       BG122005                  93 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602351665F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4449713 5',
            mRNA sequence.
ACCESSION   BG122005
VERSION     BG122005.1  GI:12615514
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10234  row: b  column: 18
            High quality sequence stop: 93.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4449713"
                     /clone_lib="NIH_MGC_90"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT       20 a      8 c     10 g     55 t
ORIGIN


  Query Match             1.7%;  Score 41;  DB 10;  Length 93;
  Best Local Similarity   76.9%;  Pred. No. 1.2e+04;
  Matches   50;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2301 ttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaa 2360
           |||| |  ||| ||||| ||||| | ||  |   |||   ||||||||||||||||||||
Db     90 TTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 31

Qy   2361 aaaaa 2365
          |||||
Db     30 AAAAA 26


RESULT  15
BE319731/c
LOCUS       BE319731                  98 bp    mRNA    linear   EST 20-DEC-2000
DEFINITION  NF020E07RT1F1051 Developing root Medicago truncatula cDNA clone
            NF020E07RT 5', mRNA sequence.
ACCESSION   BE319731
VERSION     BE319731.2  GI:11930002
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
            Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May
            ,G.D. and Paiva,N.L.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula root library
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:9193508.
            Contact: Paiva NL
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7317
            Fax: 580 221 7380
            Email: nlpaiva@noble.org
            Medicago Genome Initiative accession: MGI:S:16962
            Insert Length: 721   Std Error: 0.00
            Plate: 020  row: E  column: 07
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF020E07RT"
                     /clone_lib="Developing root"
                     /tissue_type="root"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Total RNA was extracted from
                     non-nodulated roots of plants grown in 1 mM nitrate
                     medium.  Samples were taken at four time points
                     (approximately two days, one, two and six weeks post
                     germination) representing early seedling growth to
                     nitrogen limitation."
BASE COUNT       15 a      6 c      0 g     77 t
ORIGIN


  Query Match             1.7%;  Score 41;  DB 10;  Length 98;
  Best Local Similarity   76.9%;  Pred. No. 1.2e+04;
  Matches   50;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   2301 ttttatactttctttaataaaaatagaagtatttaaacttaaaaaaaaaaaaaaaaaaaa 2360
           |||| |  ||| ||||| ||||| | ||  |   |||   ||||||||||||||||||||
Db     97 TTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 38

Qy   2361 aaaaa 2365
          |||||
Db     37 AAAAA 33


Search completed: September 24, 2002, 19:56:52
Job time: 2874 sec
```